

University of Patras
Ohio State University
Ghent University

PROCEEDINGS

4TH INTERNATIONAL
CONFERENCE OF
MODERN GREEK
DIALECTS AND
LINGUISTIC
THEORY

PATRAS 2010

EDITORS:
ANGELA RALLI, BRIAN D. JOSEPH,
MARK JANSE, & ATHANASIOS KARASIMOS,



Research on Greek Dialects: *Institutions and Projects*

On-line Proceedings of the
Forth International Conference of
**Modern Greek Dialects and
Linguistic Theory (MGDLT4)**
Chios, 11-14 June 2009

Edited by:
*ANGELA RALLI, BRIAN D. JOSEPH, MARK JANSE AND
ATHANASIOS KARASIMOS*

URL: <http://www.philology.upatras.gr/LMGD/el/index.html>

2010 University of Patras, Greece
On-line Proceedings of the Forth International Conference of
Modern Greek Dialects and Linguistic Theory (MGDLT4)
ISSN: 1792-3743

E-Proceedings Board

Editors:

Prof. Angela Ralli, University of Patras
Prof. Brian D. Joseph, Ohio State University
Prof. Mark Janse, Ghent University

Volume Editor:

Athanasios N. Karasimos, University of Patras

Organizing Committee:

Angela Ralli (chair)
Dimitrios Papazachariou
Dimitra Melissaropoulou
Eleonora Dimela
Athanasios Karasimos
Nikolaos Koutsoukos
Maria Koliopoulou

Sponsors:

University of Patras
Prefecture of Chios
Loukas Ktistakis

Contact Information:

Centre of Modern Greek Dialects
Department of Philology
University of Patras
26504 Rio, Greece
<http://www.philology-upatras.gr>
Tel/Fax: +30 2610 969375
email: ralli@upatras.gr
akarasimos@upatras.gr

ISSN: 1792-3743

Contents

Foreword

ANGELA RALLI, DIMITRIOS PAPAACHARIOU & ATHANASIOS KARASIMOS <i>Εργαστήριο Νεοελληνικών Διαλέκτων και η βάση δεδομένων GREE.D</i>	7
ΓΙΩΡΓΟΣ ΠΑΠΑΝΑΣΤΑΣΙΟΥ <i>Η έρευνα για τις νεοελληνικές διαλέκτους στο Ινστιτούτο Νεοελληνικών Σπουδών</i>	15
CHRISTINA BASSEA-BEZANTAKOU <i>Research Center for Modern Greek Dialects - Historical Dictionary</i>	18
ΙΟ ΜΑΝΟΛΕΣΣΟΥ <i>Modern Greek Dialectology and the Grammar of Medieval Greek</i>	24
ARGIRIS ARCHAKIS, SOFIA LAMPROPOULOU & DIMITRIS PAPAACHARIOU <i>Performing dialectal talk: Differentiating gender roles via direct speech representation</i>	31
SPYROS ARMOSTI <i>The perception of plosive gemination in Cypriot Greek</i>	39
MARY BALTAZANI & NINA TOPINTZI <i>The phonology and phonetics of glides in North-Western Greek dialects</i>	61
GIANNOULA GIANNOULOPOULOU <i>Diminutive suffixes in modern greek dialects</i>	75
KLEANTHES K. GROHMANN & ELENA PAPAADOPOULOU <i>Cypriot anomalies in wh-in situ structures</i>	81
NIKOS LIOSIS <i>Counterfactuality in the Tsakonian dialect: a contribution to the history of 'ήθελα' and 'ήμουν'</i>	111
ANASTASSIA LOUKINA <i>Towards the acoustic analysis of lateral consonants in Modern Greek dialects: a preliminary study</i>	124
DIMITRIS MICHELIOUDAKIS & IOANNA SITARIDOU <i>Syntactic micro-variation in Pontic: Dative constructions</i>	137
ALEXANDER PAVLENKO & GALINA PAVLENKO <i>The impact of Greek on the northern Azov varieties of Russian</i>	150
NATALIA PAVLOU	

<i>MBU! On wh-objects and true adjuncts of Cypriot Greek</i>	156
ANTONIO ROMANO, ΦΡΑΝΤΣΕΣΚΑ ΠΑΠΑΣΠΥΡΟΥ & ΠΑΥΛΟ ΜΑΪΡΑΝΟ <i>Ορισμένες σκέψεις σχετικά με την προσωδία του Γκρίκου</i>	179
ANNETTE TEFFETELLER <i>Object clitics in the Modern Greek dialects of Asia Minor: diachronic and dialectal variation in the encoding of argument</i>	187
NINA TOPINTZI, KATERINA NICOLAIDIS & ELENI TSIARTSIONI <i>Inter-dialectal Insights into Greek Rhythm: The Case of Standard Modern Greek vs. Kozani Greek</i>	196
MARINA TZAKOSTA <i>“The importance of being voiced”: cluster formation in dialectal variants of Greek</i>	212
ΣΥΜΕΩΝ ΤΣΟΛΑΚΙΔΗΣ <i>Αντιγεγονοτικές περιφράσεις με το ήθελα+απρωφ. στις νεοελληνικές διαλέκτους</i>	223
ELEONORA DIMELA & ANGELA RALLI <i>moro, akro, sjo: Prefixes or Compound Constituents?</i>	230
DIMITRA MELISSAROPOULOU <i>Loan verb adaptation in Greek dialectal variation: A first approach</i>	243
MARIA KAL TSA & IOANNA SITARIDOU <i>Topicalisation in Pontic Greek</i>	259
STERGIOS CHATZIKYRIAKIDIS <i>Clitic climbing in Grecia Salentina Greek: A dynamic account</i>	280
PANAYIOTIS A. PAPPAS <i>A new sociolinguistic variable in Cypriot Greek</i>	305
NIKOLAOS PANTELIDIS <i>The Reshaping of the Mediopassive Endings: Evidence from Modern Greek Varieties</i>	315

Foreword

The *Fourth International Conference on Modern Greek Dialects and Linguistic Theory (MGDLT4)* was held at the Homerion Institute of Chios, Greece (June 11-14, 2009). It was hosted by the Prefecture of Chios, with the support of the University of Patras (Research Committee), and Loukas Ktistakis Shipping Company. It was chaired by Prof. Angela Ralli (University of Patras), Prof. Brian D. Joseph (Ohio State University), and Prof. Mark Janse (Ghent University).

The conference brought together experts working on both linguistics and the dialects of Modern Greek, in a variety of topics and orientations. In the first day, there was a special workshop about *Research on Greek Dialects, Institutes, and Projects relative to Dialectology*. The speakers were: Dr. Christina Bassea-Bezantakou (Research Center for Modern Greek Dialects - Historical Dictionary, Academy of Athens), Dr. George Papanastasiou (Institute of Modern Greek Studies - Manolis Triantaphyllidis Foundation), Dr. Io Manolessou (Grammar of Medieval Greek, University of Cambridge), Prof. Panagiotis Kontos (University of Athens), Prof. Angela Ralli, Dr. Dimitris Papazachariou, and Athanasios Karasimos (University of Patras, Centre of Modern Greek Dialects).

The Scientific Committee wishes to express its gratitude to the invited speakers, who so promptly responded to its call, namely, Cleo Condoravdi (Stanford University), and Anna-Maria Di Sciullo (UQAM Canada). It would also like to thank the other speakers for their participation.

A special thank is due to Prof. Stavros Koubias, Rector of the University of Patras, Mr. Loukas Ktistakis, as well as the Research Committee of the University of Patras, and the Prefecture of Chios, whose generous support, moral and financial, made the organization of the Conference possible.

Last, but not least, the Scientific Committee is particularly grateful to the members of the Organizing Committee, Dimitra Melissaropoulou, Eleonora Dimela, Athanasios Karasimos, Nikolaos Koutsoukos, Maria Koliopoulou, and Nikos Angelopoulos, for their most valuable help before and during the conference.

The Permanent Scientific Committee

Angela Ralli
Brian D. Joseph
Mark Janse

Εργαστήριο Νεοελληνικών Διαλέκτων και η βάση δεδομένων GREED

ΑΓΓΕΛΙΚΗ ΡΑΛΛΗ
ralli@upatras.gr

ΔΗΜΗΤΡΙΟΣ ΠΑΠΑΖΑΧΑΡΙΟΥ
Πανεπιστήμιο Πατρών
papaz@upatras.gr

ΑΘΑΝΑΣΙΟΣ ΚΑΡΑΣΙΜΟΣ
akarasimos@upatras.gr

1. Εισαγωγή

Η Νέα Ελληνική είναι πλούσια σε διαλεκτικές ποικιλίες, οι οποίες χρησιμοποιούνται στον καθημερινό λόγο, ενώ υπάρχουν και κάποιες γλωσσικές ποικιλίες που περιορίζονται σε συγκεκριμένες ομάδες πρεσβύτερων/ γερόντων και αντιμετωπίζουν το φάσμα της εξαφάνισης και εξάλειψης (Trudgill 1998, Κοντοσόπουλος 2001).

Εντούτοις, οι διαλεκτικές ποικιλίες μελετήθηκαν ελάχιστα, αν και περιέχουν αξιοσημείωτη εμφάνιση φαινομένων για τη γλωσσολογική ανάλυση. Αυτό το διαλεκτικό μωσαϊκό οφείλεται σε μεγάλο βαθμό σε συγκεκριμένες ιστορικές, πολιτικές και κοινωνικές συνθήκες και περιστάσεις που χαρακτήθηκαν στην Ιστορία του Νεότερου Ελληνικού Κράτους, που απελευθερώθηκε από την Οθωμανική Αυτοκρατορία στις αρχές του 19^{ου} αιώνα και κατά την αρχική του σύσταση περιλάμβανε τις γεωγραφικές περιοχές της Πελοποννήσου, της Στερεάς Ελλάδας και κάποιων νησιών. Έως τότε, διάφορες ομάδες της σύγχρονης Ελλάδας έκαναν εσωτερική μετανάστευση στο νεοσύστατο κράτος (π.χ. από Κρήτη, Μακεδονία και Δωδεκάνησα), ενώ παράλληλα ένας σημαντικός αριθμός Ελλήνων διαλεκτόφωνων προσφύγων μετακινήθηκαν από την Τουρκία (Μικρά Ασία και Πόντος) στην Ελλάδα, με το πέρας της Μικρασιατικής καταστροφής το 1922 και την ανταλλαγή πληθυσμών.

Σήμερα, η Κοινή Νέα Ελληνική είναι κυρίως βασισμένη στην Πελοποννησιακή διάλεκτο, ενώ οι διάλεκτοι από τα υπόλοιπα γεωγραφικά διαμερίσματα εντός και εκτός Ελλάδος δημιουργούν ένα ιδιαίτερο, ξεχωριστό και ποικιλόχρωμο γλωσσικό μωσαϊκό, οι οποίες χρήζουν άμεσα να περιγραφούν, να αναλυθούν και να διατηρηθούν, προτού αυτές εξαλείψουν παντελώς.

Εντούτοις, προς τη συγκεκριμένη κατεύθυνση δεν έχουν γίνει σοβαρά και συστηματικά βήματα έρευνας. Στην Ελλάδα υπάρχει από το 1908 ένα εθνικό ερευνητικό κέντρο στην Ακαδημία Αθηνών, το οποίο ενδιαφέρεται για γραπτά και προφορικά διαλεκτικά δεδομένα, αλλά τα διαλεκτικά δεδομένα δεν είναι ψηφιοποιημένα, τα περισσότερα είναι αδημοσίευτα με αυξημένες δυσκολίες πρόσβασης για τους εξωτερικούς ερευνητές. Μη-ψηφιοποιημένα διαλεκτικά δεδομένα εντοπίζονται παράλληλα σε συγκεκριμένους συλλόγους και οργανισμούς από πρόσφυγες από κάθε γωνιά της Ελλάδος, όπως για παράδειγμα το Ιστορικό Αρχείο των Μικρασιατών Ελλήνων στη Θεσσαλονίκη, το κέντρο Μικρασιατικών σπουδών, η Ένωση Ποντίων στην Παναγία Σουμελά Ημαθίας, αλλά έχουν συλλεχθεί κυρίως με ιστορικά κριτήρια και στόχους και φυσικά δεν έχουν ταξινομηθεί και κατηγοριοποιηθεί συστηματικά.

Η πρώτη συστηματική προσπάθεια ψηφιοποίησης, καταλογογράφησης και κωδικοποίησης διαλεκτικών δεδομένων έγινε από το Εργαστήριο Νεοελληνικών Διαλέκτων του Πανεπιστημίου Πατρών με την υλοποίηση της ηλεκτρονικής βάσης GREED, η οποία περιέχει γλωσσολογικά και μετα-γλωσσολογικά *corpora*. Αυτά τα δεδομένα συλλέχθηκαν από έρευνες πεδίου, όπου καταγράφηκαν δεδομένα φυσικής και αυθόρμητης ομιλίας με στόχο το σχηματισμό μιας αντιπροσωπευτικής εικόνας της γλωσσολογικής κατάστασης συγκεκριμένων γεωγραφικών και κοινωνικών περιοχών της Ελλάδος. Παράλληλα, γίνεται προσπάθεια συλλογής και διαλεκτικών χειρογράφων και διαφόρων κειμένων, βιβλίων, έντυπων συλλογών, ώστε να δημιουργήσουμε ένα

ψηφιοποιημένο σώμα κειμένων, ωστόσο ο τελευταίος στόχος αποτελεί μακροχρόνια προσπάθεια και έμμεση προτεραιότητα. Φιλοδοξία μας είναι η βάση GREED να αποτελεί πολύτιμο αρωγό για τη μελλοντική έρευνα της κατηγοριοποίησης και οργάνωσης των διαφόρων γλωσσολογικών φαινομένων – φωνολογικά, μορφολογικά, κοινωνιογλωσσολογικά κτλ. – που εντοπίζονται διαδialeκτικά. Επομένως, θα διευκολύνει αισθητά τις δημοσιεύσεις και εκδόσεις γλωσσάριων, λεξικών και γραμματικών των διαφόρων διαλέκτων της Νέας Ελληνικής.

2. GREED Corpus και συλλογή δεδομένων

Ο θεμέλιος λίθος για την ανάπτυξη της ηλεκτρονικής βάσης GREED αποτέλεσαν διάφορα ερευνητικά προγράμματα που αποσκοπούσαν στη διατήρηση συγκεκριμένων διαλέκτων: “*Grico: Dialect spoken in the area of Salento, South Italy*” (Interreg II, Ευρωπαϊκή Ένωση, σύνολο 55 ωρών, συντονίστρια Αγγελική Ράλλη).

“*Διαλεκτικές ποικιλίες της Ανατολικής Λέσβου. Σύγκριση με την μικρασιατική διάλεκτο των Κυδωνίων και Μοσχονησίων*” (Υπουργείο Παιδείας, σύνολο 45 ωρών, συντονίστρια Αγγελική Ράλλη).

“*Η μικρασιάτικη διάλεκτος των Κυδωνίων και Μοσχονησίων*” (Υπουργείο Αιγαίου και Υπουργείο Παιδείας, σύνολο 112 ώρες, συντονίστρια Αγγελική Ράλλη).

“*Cappadocian*”. Endangered Languages and Documentation Programme. University of London SOAS, σύνολο 40 ωρών, συντονιστές Mark Janse, Αγγελική Ράλλη και Δημήτρης Παπαζαχαρίου).

“*Διαλεκτική ποικιλία Πάτρας*” (Πανεπιστήμιο Πατρών, σύνολο 100 ωρών, συντονιστής Δημήτρης Παπαζαχαρίου).

“*Η διάλεκτος της Αγίας Παρασκευής Λέσβου*” (Δήμος Αγίας Παρασκευής, σύνολο 40 ωρών, συντονίστρια Αγγελική Ράλλη)

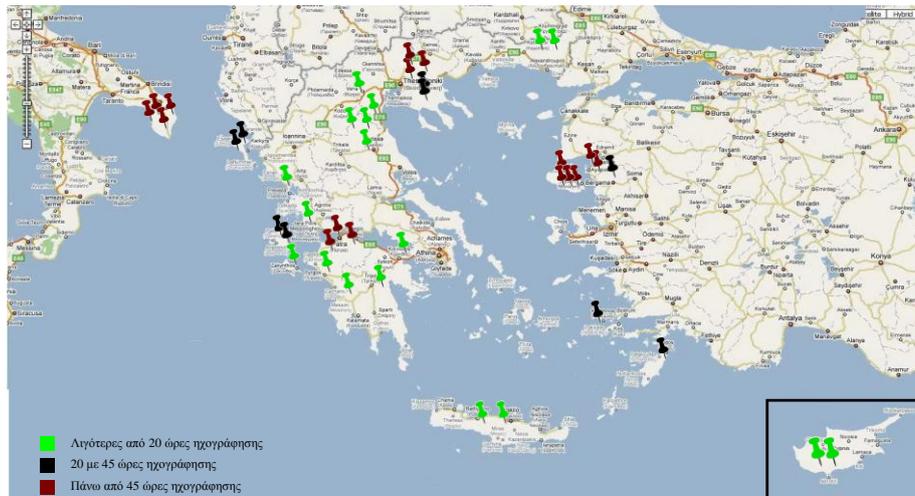
“*Τουρκοκρητικά Μικράς Ασίας*” (Υπουργείο Εξωτερικών, σύνολο 32 ωρών, συντονίστριες Αγγελική Ράλλη και ΧΧ)

“*Από το γλωσσικό ιδίωμα των Μεγάρων στο γλωσσικό ιδίωμα της Παλαιάς Αθήνας*” (Ίδρυμα Λεβέντη και Δήμος Μεγαρέων, σύνολο 44 ωρών, συντονίστριες Αγγελική Ράλλη και Αγγελική Σύρκου)

Παράλληλα η συλλογή υλικού γίνεται στα πλαίσια μαθημάτων, διπλωματικών εργασιών και διδακτορικών διατριβών που προσθέτουν στη βάση σημαντικό υλικό. Ο ακόλουθος πίνακας δίνει μια κατατοπιστική εικόνα του συνολικού υλικού της βάσης:

Διαλεκτική περιοχή	Ώρες	Ποσοστό	Ομιλητές	Ποσοστό
Καππαδοκικά	41 Ώρες	8%	82 Ομιλητές	12,77%
Μικρά Ασία	105 Ώρες	21,00%	78 Ομιλητές	12,14%
Κύπρος	2,5 Ώρες	0,50%	12 Ομιλητές	1,89%
Δωδεκάνησα	9,5 Ώρες	2%	13 Ομιλητές	2,02%
Ήπειρος	12 Ώρες	2,20%	17 Ομιλητές	2,60%
Επτάνησα	15 Ώρες	3,00%	33 Ομιλητές	5,10%
Μακεδονία	9 Ώρες	1,60%	16 Ομιλητές	2,50%
Λέσβος	128 Ώρες	25,30%	80 Ομιλητές	12,46%
Κάτω Ιταλία	55 Ώρες	11,00%	68 Ομιλητές	10,60%
Στερεά Ελλάδα	12 Ώρες	2,20%	21 Ομιλητές	3,27%
Θεσσαλία	8 Ώρες	2%	16 Ομιλητές	2,50%
Θράκη	8 Ώρες	2%	6 Ομιλητές	1%
Πελοπόννησος			200	
	100 Ώρες	20%	Ομιλητές	31,15%
Σύνολο	505 Ώρες	100,0 %	642	
			Ομιλητές	100,0%

Πίνακας 1: Συνολική στατιστική της ηλεκτρονικής βάσης



Εικόνα 1: Οι γεωγραφικοί τόποι, όπου πραγματοποιήθηκαν διαλεκτικές ηχογραφήσεις από το Εργαστήριο Νεοελληνικών Διαλέκτων

Ουσιαστικά υπάρχουν δύο διαφορετικοί τύποι του προφορικού υλικού: α) ηχογραφήσεις από αυθόρμητο προφορικό λόγο από συναντήσεις και β) στοχευμένες συνεντεύξεις για την εξαγωγή συγκεκριμένων γλωσσολογικών πληροφοριών από προφορικό υλικό. Η προσπάθεια αναμόχλευσης προσωπικών διηγήσεων και κυρίως παλαιότερων ιστοριών ήταν ηελεημένη επιλογή – στις πλειονότητα των ερευνητικών προγραμμάτων – για τη διαφύλαξη υλικού πολιτισμικής κληρονομιάς ταυτόχρονα με τη συλλογή των γλωσσικών δεδομένων. Πιο συγκεκριμένα, οι ηχογραφήσεις έγιναν από ερευνητές πεδίου που είχαν αποκτήσει κάποιες κοινωνικές σχέσεις και επαφές με την υπό διερεύνηση κοινότητα και τους πληροφορητές συγκεκριμένα, ή κυρίως με τη συνδρομή της φυσικής παρουσίας ενδιαμέσου, δηλαδή ενός μέλους της τοπικής κοινότητας ή άτομο που διατηρεί στενές επαφές με τους πληροφορητές (φίλος φίλου, συγγενής συγγενή, γείτονας). Για παράδειγμα, μιλώντας για τις προσωπικές εμπειρίες και δυσκολίες από δύσκολες περιόδους της ελληνικής ιστορίας ήταν πιο αποτελεσματικό για να τους κάνουμε να ανοιχτούν συναισθηματικά, να αισθανθούν άνετα και να μπορέσουν να εκφραστούν ελεύθερα μιλώντας διαλεκτικά και να καταφέρουν να αφαιρέσουν από το μυαλό την ιδέα της συνέντευξης και να αισθανθούν ότι βρίσκονται σε μια καθημερινή στιγμή. Σύμφωνα με τις αρχές της Μεθοδολογίας της Έρευνας, αυτή η μέθοδος παρέχει σημαντικές πληροφορίες για την προφορική ιστορία και τα γλωσσικά δεδομένα και αυξάνει σημαντικά τις πιθανότητες για συλλογή αυθόρμητου λόγου και περιορισμό του φαινομένου της προσποίησης.

Στην GREE.D, οι διάλεκτοι είναι καταχωρημένες γεωγραφικά (καθότι αυτή πληροφορία θα βοηθήσει το διαλεκτικό χάρτη μελλοντικά) και οι πληροφορίες σχετικά με τα μεταδεδομένα είναι δομημένες σε επτά βασικές κατηγορίες: Ιδιότητες Αρχείων, Διάλεκτος, Ερευνητικό πρόγραμμα, Τεχνικές πληροφορίες, Επικοινωνιακή κατάσταση, Πληροφορητές, Γλωσσολογικά δεδομένα. Αυτές οι βασικές κατηγορίες που χρησιμοποιήθηκαν και για τον χαρακτηρισμό όλου του προφορικού υλικού, έχουν πολλές υποκατηγορίες που παρέχουν πολλές επιλογές για τη δημιουργία μιας προχωρημένης μηχανής αναζήτησης. Στα σχήματα που ακολουθούν δίνονται δείγματα από δύο ομάδες μεταδεδομένων από τη συλλογή προφορικού υλικού από τις Νεοελληνικές διαλέκτους.

Αν και η δημιουργία της βάσης εξακολουθεί να είναι υπό δημιουργία, η GREE.D περιέχει πάνω από 460 ώρες προφορικού υλικού, συνοδευμένο από μεταδεδομένα και 40 ώρες του υλικού έχει ήδη απομαγνητοφωνηθεί συνοδευμένο από πρωτόκολλο χαρτογράφησης αρχείων.

Παράλληλα έχει ξεκινήσει η ψηφιοποίηση διαφόρων χειρογράφων και σπανίων βιβλίων, ώστε σύντομα η GREED να διαθέτει και την αντίστοιχη πλατφόρμα για την αξιοποίηση του γραπτού διαλεκτικού υλικού. Στο πλαίσιο του ερευνητικού προγράμματος έγινε μια οργανωμένη προσπάθεια ψηφιοποίησης χειρογράφων, κυρίως νομικής και συμβολαιογραφικής φύσεως που περιείχαν μεταξύ άλλων και διαλεκτικό υλικό. Τα 1500 και πλέον χειρόγραφα ψηφιοποιήθηκαν και βρίσκονται στη διαδικασία χαρακτηρισμού τους από επιλεγμένες πληροφορίες μεταδεδομένων. Είναι ενδεικτικό ότι ηλεκτρονικές βιβλιοθήκες χειρογράφων στο διαδίκτυο συνοδεύονται πάντα από πληροφορίες περιγραφής του χειρογράφου.¹

3. Είδος αρχείων (αρχεία ήχου, μεταγραφές, πρωτόκολλα χαρτογράφησης και ψηφιακά χειρόγραφα)

Τα δεδομένα των ηχογραφήσεων των διαλέκτων συλλέχθηκαν με τη χρήση επαγγελματικών ψηφιακών κασετοφώνων Marantz. Η επιλογή των επαγγελματικών ψηφιακών συσκευών ελήφθη με βάση τις διεθνείς προδιαγραφές για ποιοτικές ηχογραφήσεις με τις ελάχιστες δυνατές απώλειες.

Οι πληροφορητές συνήθως ηχογραφήθηκαν κατά ζεύγη ή κατά μόνες με τη συνδρομή του ενδιαμέσου και ο μέσος χρόνος ηχογράφησης είναι περίπου στα εξήντα λεπτά. Όπως και στα πιο πρόσφατα ερευνητικά προγράμματα, οι ηχογραφήσεις πραγματοποιήθηκαν με ψηφιακές συσκευές εγγραφής (η επαγγελματική σειρά της Marantz), που εγγράφει τις συνομιλίες σε ασυμπιεστη μορφή αρχείου .wav και ελαχιστοποιεί την οποιαδήποτε διαδικασία ψηφιοποίησης των ηχητικών αρχείων. Παράλληλα, οι συγκεκριμένες συνομιλίες καταγράφονται στερεοφωνικά – σε αριστερό και δεξί κανάλι – με τη χρήση δύο μικροφώνων, ώστε να αντιστοιχείται ένα κανάλι ανά πληροφορητή, εφόσον είναι δυνατόν. Με αυτό τον τρόπο, καταφέραμε να μειώσουμε το περιβαλλοντικό θόρυβο (περίπου 40 db) για να επιτύχουμε την μέγιστη δυνατή ποιότητα εγγραφής και την ίδια στιγμή να μειώσουμε στο ελάχιστο το προβληματικό φαινόμενο της επικάλυψης, όταν δύο ομιλητές μιλάνε την ίδια χρονική στιγμή ή διακόπτει ο ένας τον άλλον.

Να σημειωθεί ότι τα ηχητικά αρχεία εισάγονται σε υπολογιστή συνδεδεμένο με βάση δεδομένων χωρίς καμία υποβάθμιση ποιότητας και αποθηκεύονται για λόγους ασφαλείας σε ένα σύστημα αποθήκευσης NAS για υψηλότερη ασφάλεια. Επίσης η εισαγωγή των ηχητικών αρχείων των διαλέκτων. Τυπικοί στόχοι επεξεργασίας συμπεριλαμβάνουν την ορθή ονοματοδοσία, διαχωρισμό καναλιών, αφαίρεση προσωπικών πληροφοριών, ενίσχυση των χαμηλής έντασης ηχογραφήσεων, μείωση του θορύβου και καθαρισμός του σήματος από έντονους μικροφωνισμούς.

Επομένως τα ηλεκτρονικά αρχεία των Νεοελληνικών διαλέκτων στην ηλεκτρονική βάση GREED είναι τα ακόλουθα:

(α.) Ψηφιακά Αρχεία ήχου: ηχογραφήσεις φυσικού διαλεκτικού λόγου σε μορφή στέρεο, καθώς και μονοκαναλικός διαχωρισμός.

(β.) Αρχεία περιγραφής των ηχογραφήσεων: (i.) μεταγραφές ομιλίας (εναλλαγές διαλόγου, απομαγνητοφώνηση ορθογραφική, φωνολογική (σπάνια) και μορφολογική σήμανση), (ii.) πρωτόκολλο χαρτογράφησης ηχητικού αρχείου (ανά δύο λεπτά χαρακτηρισμός αρχείου με συγκεκριμένα κριτήρια

¹ Ενδεικτικά η Schoenberg Database of Manuscripts

(<http://dla.library.upenn.edu/cocoon/dla/schoenberg/index.html>), η National Mission for Manuscripts (<http://www.namami.org/index.htm>), η Leeds Verse Database

(<http://www.leeds.ac.uk/library/spcoll/bcmsgv/intro.htm>), η International Dunhuang Project: The

Silk Road Online (<http://idp.bl.uk/>), η Medieval and Early Modern Manuscripts Collection:

Database and Digital Images (<http://research.hrc.utexas.edu/pubmnem/>), η Old English

Manuscript Database

(<http://www8.georgetown.edu/departments/medieval/labyrinth/subjects/mss/oe/oldeng.html>)

μεταξύ άλλων.

(γ.) Κείμενα και χειρόγραφα: κείμενα που έχουν γραφτεί πρωταρχικώς στη διάλεκτο. Εκτός από τα ψηφιακά αρχεία ήχου, μια ικανοποιητική βάση δεδομένων πρέπει να εσωκλείει και μεταγραφές – απομαγνητοφωνήσεις των αρχείων. Υπάρχει μια μεγάλη συζήτηση από τους ερευνητές βάσεων δεδομένων για το ποιος είναι ο πλέον κατάλληλος τρόπος μεταγραφής των ηχητικών αρχείων (*φωνητικός*, *φωνολογικός* ή *ορθογραφικός*). Συμφωνώντας με τους Durand & Eriksson (2007) και τους Anderwald & Wagner (2007: 42-43) υποστηρίζουμε ότι τα μειονεκτήματα της φωνολογικής και φωνητικής απομαγνητοφώνησης είναι τέτοιας φύσεως για τα ελληνικά που προτιμήσαμε την ορθογραφική μεταγραφή των προφορικών συνομιλιών. Η επιλογή μας επηρεάστηκε σημαντικά από την προοπτική εκμετάλλευσης του διαλεκτικού υλικού για μορφολογικούς αναλυτές με τη χρήση του απομαγνητοφωνημένου υλικού για μορφολογικούς και λεξικογραφικούς σκοπούς. Παράλληλα κατά την απομαγνητοφώνηση βασιστήκαμε στις κωδικοποιήσεις της Ανάλυσης Λόγου αναφορικά με τις εναλλαγές διαλόγου, διακοπές, επικαλύψεις, παύσεις, επιμηκύνσεις, γρήγορος ή αργός ρυθμός ομιλίας, ένταση και χαμηλόφωνη ομιλία, είναι τα διάφορα μεταγλωσσικά φαινόμενα που μπορούν να επηρεάσουν φωνολογικά φαινόμενα και σημειώνονται κατά την απομαγνητοφώνηση και χαρτογράφηση του αρχείου.

Η ορθογραφική μεταγραφή δίνει τη δυνατότητα για πιο απρόσκοπτη διερεύνηση των μορφοσυντακτικών χαρακτηριστικών και κοινωνιογλωσσολογικών φαινομένων, αλλά υπάρχουν εμφανή προβλήματα που αφορούν ζητήματα τεχνικής φύσεως, όπως για παράδειγμα, πως θα λειτουργήσει η φωνητική κωδικοποίηση σε λογισμικά όπως το Praat και το E-Lan.

Τέλος, μόνο η ορθογραφική μεταγραφή των δεδομένων θα καλύψει τις υπάρχουσες απαιτήσεις της βάσης: στόχος ενός ολοκληρωμένου corpus πρέπει να είναι η δυνατότητα να είναι μηχανικά-αναγνώσιμο (*machine-readable*), να επιτρέπει την εύκολη και γρήγορη διαχείριση αναζήτησης με διάφορα εργαλεία και το πλέον σημαντικό να συγκρίνεται με άλλα σώματα κειμένων όσον αφορά την απλότητα και την ευχρηστία. Επιπροσθέτως, η ορθογραφική μεταγραφή θα μας επιτρέψει να συγκρίνουμε τα δεδομένα με αντίστοιχα άλλων γραπτών και προφορικών συλλογών και μας επιτρέπουν να κάνουμε συγκρίσεις ανάμεσα σε διαφορετικούς ομιλητές, διαφορετικές διαλέκτους και διαλεκτικές περιοχές και διαφορετικά corpora.

Παρόλο που οι συνεντεύξεις είναι άμεσα προσβάσιμες λόγω της ηλεκτρονικής τους μορφής [ο κάθε ερευνητής μπορεί να έχει άμεση πρόσβαση στο αρχείο που επιθυμεί για ανάλυση, ακόμα και στην στερεοφωνική του μορφή], η απουσία φωνολογικής απομαγνητοφώνησης αποτρέπει την γρήγορη και ευρεία φωνολογική ανάλυση χωρίς τη χρήση των ηχητικών αρχείων. Όλα τα απομαγνητοφωνημένα αρχεία έχουν καταγραφεί και σε αρκετά σημεία φωνολογικά φαινόμενα έχουν χαρτογραφηθεί από την απομαγνητοφώνηση χωρίς την άμεση σύνδεση με τα ηχητικά αρχεία. Ελπίζουμε μελλοντικά πως η ηλεκτρονική βάση θα παρέχει την επιθυμητή ευθυγράμμιση ήχου και κειμένου, όπως στο Necte (βλ. Allen *et al.* 2007) και στο ONZE² (βλ. Gordon *et al.* 2007): προς το παρόν η ευθυγράμμιση επιτυγχάνεται μόνο μέσω του E-Lan και του Praat.

Για να καλυφθούν κάποια κενά της ορθογραφικής μεταγραφής, αλλά κυρίως για την δυνατότητα μιας γρήγορης χαρτογράφησης και «ακτινογραφίας» ενός ηχητικού αρχείου παρέχεται για αρκετές περιπτώσεις των διαλεκτικών δεδομένων το πρωτόκολλο χαρτογράφησης. Ανά δύο λεπτά χαρακτηρίζεται το αρχείο με βάση κάποια κριτήρια τεχνικά και περιγραφικά, όπως ποιότητα ηχογράφησης, ύπαρξη θορύβων, αριθμός ομιλητών, καθώς και με γλωσσολογικά κριτήρια, όπως καταγραφή ή σήμανση ενδιαφερόντων γλωσσικών φαινομένων πάσης φύσεως (π.χ. σήμανση για αλλόμορφα, για ασυνήθιστο επιτονισμό, για συντακτικούς περιορισμούς κλπ.).

² <http://www.lacl.canterbury.ac.nz/onze/news.html>

4. Διαχείριση και ιστοσελίδα

Οι απαιτήσεις για τη βάση δεδομένων είναι για ένα σύστημα που να μπορεί να παρέχει πρόσβαση στα διαλεκτικά δεδομένα μέσω μιας κοινής διεπιφάνειας. Απαιτητικοί έλεγχοι πιστότητας πρωτοκόλλων και λοιποί κανόνες σχετικά με συνοχή και ασφάλεια των δεδομένων αποτελούν βασικές προϋποθέσεις. Κις και πρωταρχικός στόχος είναι η υλοποίηση ενός εργαλείου βάσης δεδομένων που να είναι εύχρηστο, πολυχρηστικό και ανοιχτό για τη γλωσσολογική κοινότητα για αρκετό καιρό, δημιουργήθηκε μια διαδραστική ιστοσελίδα (έχοντας ως οδηγό τα ISCC χαρακτηριστικά, Dipper et al. 2007) με στόχο να μπορεί να αλληλεπιδρά με άλλα λογισμικά επεξεργασίας, όπως Praat. Το σύστημά μας υποστηρίζει ελληνικούς και λατινικούς χαρακτήρες. Το περιβάλλον εργασίας των χρηστών που παρέχεται στους ερευνητές είναι γρήγορο και εύκολο στη χρήση· επομένως ο χρόνος εκπαίδευσης είναι μειωμένος.

Η αρχιτεκτονική δομή της βάσης είναι χτισμένη πάνω σε τέσσερα αντικείμενα. Όλα τα αντικείμενά (*Metadata*, *Metadatatetails*, *mdListValues* [προ-εισαγμένες τιμές] και *FileAttribs* [πίνακας με όλα τα αρχεία]) είναι συνδεδεμένα αναμεταξύ τους με μια σχέση 'ένα προς πολλά', για παράδειγμα η τιμή 'dialect name' του *Metadatatetails* είναι συνδεδεμένη με τις τιμές 'Ποντιακά', 'Λεσβιακά', 'Κυπριακά' μεταξύ άλλων τιμών από το *mdListValues*. Το σύστημά είναι βασισμένο σε αρχιτεκτονική client-server (apache server), η οποία συσχετίζεται με μια συσχετιστική βάση δεδομένων τύπου MySQL. Όλες οι σελίδες είναι χτισμένες πάνω σε φόρμες template και επεξεργάζονται τα δεδομένα χρησιμοποιώντας μικρούς κώδικες σε PHP γλώσσα. Οι χρήστες έχουν πρόσβαση στα δεδομένα μέσω μιας PHP διεπιφάνειας με τη χρήση του HTML πρωτοκόλλου. Ένας σημαντικός λόγος επιλογής ενός client/ server δικτύου είναι επειδή επιτρέπει την πρόσβαση στη βάση δεδομένων την ίδια στιγμή και στα αρχεία που είναι αποθηκευμένα στον server.

Το βασισμένο στο διαδίκτυο σύστημα μας ακολουθεί τις αρχές ενός client/ server μοντέλου σχετικά με την προσκόμιση πληροφορίας των αρχείων. Βασισμένο σε ένα τέτοιο μοντέλο ο client υπολογιστής είναι συνδεδεμένος με τον server υπολογιστή, ο οποίος περιέχει τις πληροφορίες και φυσικά ο client υπολογιστής εξαρτάται άμεσα από τον server για την απόκτηση των απαραίτητων πληροφοριών. Βασισμένο στη δικτυακή τεχνολογία, είναι ανοιχτό για οποιοδήποτε λειτουργικό σύστημα που έχει φυλλομετρητή διαδικτύου (web browser). Για την ώρα, για τη διαφύλαξη της σταθερότητας του συστήματος, οι χρήστες μπορούν να ανεβάσουν αρχεία, αλλά οι τιμές των μεταδεδομένων πρέπει να εισαχθούν από τον διαχειριστή του συστήματος έπειτα από αίτηση του χρήστη. Στην παρούσα φάση της υλοποίησης, δουλεύουμε σε μια παραλλαγμένη TEI (Text Encoding Initiative) έκδοση για τα δεδομένα. Επιπλέον, το σύστημα παράγει αναφορές καταγραφής αλλαγών και προβλημάτων αυτόματα, ώστε να είναι δυνατή η γρήγορη εύρεση του προβλήματος, για παράδειγμα όταν ο διαμοιραστής αποτυγχάνει να αναβαθμίσει τις φόρμες των απαραίτητων μεταδεδομένων μέσα σε περιορισμένο χρονικό διάστημα (30 δευτερόλεπτα).

5. Εργαλεία ανάλυσης των Νεοελληνικών Διαλέκτων

Όπως αναφέραμε σε προηγούμενη ενότητα η βάση δεδομένων συνοδεύεται εκτός από τα ηχητικά αρχεία και από τα αντίστοιχα αρχεία μεταγραφής, για όσα αρχεία ήχου έχουν πραγματοποιηθεί. Η επιλογή συνοδευτικού λογισμικού δεν είναι εύκολη υπόθεση, αποτελεί αναπόσπαστο κομμάτι μιας καλής βάσης προφορικών δεδομένων και τα λογισμικά πρέπει να πληρούν βασικά κριτήρια³:

- (1) Να είναι λογισμικά ανοιχτού κώδικα και ελεύθερα ως προς τη χρήση
- (2) Να παρέχει μεγάλο εύρος σχεδιαστικών παραμέτρων
- (3) Να υποστηρίζει αρχεία από διαφορετικά λογισμικά που χρησιμοποιούνται για τον σχολιασμό αρχείων σε διαφορετικά γλωσσολογικά επίπεδα
- (4) Να επιτρέπει την χρήση πιθανών add-ons και plug-ins

³ Για αυτό το λόγο επιλέχθηκαν τα λογισμικά Praat (μαζί με το Akustyk) και το ELAN.

- (5) Να προσφέρεται συνεχής υποστήριξη από τους προγραμματιστές/ παραγωγούς του λογισμικού
(6) Να είναι πολυγλωσσικό ή τουλάχιστον σε αγγλική έκδοση και να επιτρέπει τη χρήση του Unicode πρωτοκόλλου

6. Μελλοντικά σχέδια

Η ηλεκτρονική διαλεκτική βάση GREE.D και η συλλογή υλικού από τις Νεοελληνικές διαλέκτους είναι έρευνα υπό εξέλιξη. Είναι στις επιθυμίες και στα σχέδια μας να παρέχουμε μια ολοκληρωμένη μορφή της βάσης, η οποία θα είναι ανοιχτή για όλη την ακαδημαϊκή – και όχι μόνο – κοινότητα. Σεβόμενοι τα μελλοντικά μας σχέδια για την ηλεκτρονική βάση διαλεκτικών δεδομένων, τα ακόλουθα σημεία θεωρούμε ότι οφείλουμε να τα υπογραμμίσουμε:

[Τεχνικά] Κατά την διάρκεια της έρευνας για τις Νεοελληνικές διαλέκτους, αναβαθμίσαμε σημαντικό την διεπιφάνεια επίδρασης του χρήστη με ένα εύκολο στη χρήση web περιβάλλον, όπου δεν απαιτείται η χρήση κανενός λογισμικού από τον χρήστη. Έχουμε τη δυνατότητα να παρέχουμε μια πληθώρα κατανοητών και κατατοπιστικών κοινωνιογλωσσολογικών μεταδεδομένων, όπως και συμπληρωματικές πληροφορίες για τα ηχητικά αρχεία. Εντούτοις, πρέπει να παρέχουμε κωδικοποιημένες πληροφορίες και μεταδεδομένα για τα ψηφιακά δεδομένα, τα οποία δεν έχουν καταχωρηθεί και καταλογογραφηθεί με ενιαίο τρόπο. Η δική μας έκδοση βρίσκεται σε στάδιο δοκιμής και αναβάθμισης, αλλά έχει αποδειχθεί μέχρι στιγμής αρκετά γρήγορη και φιλική προς τον χρήστη.

[Τεχνικά] Δημιουργούμε έναν πιο αναπτυγμένο σύστημα αναζήτησης με κριτήρια βασισμένα στα μεταδεδομένα. Στοχεύουμε να κάνουμε τη βάση πιο γρήγορη, χωρίς προβλήματα και με σταθερότητα κώδικα.

[Τεχνικά] Να ελέγξουμε τα υπάρχοντα αρχεία μεταγραφής και απομαγνητοφώνησης και να συνεχίσουμε την μεταγραφή των υπόλοιπων διαλεκτικών προφορικών αρχείων.

[Τεχνικά] Έναρξη ευρύτερων φωνολογικών/ φωνητικών μεταγραφών που να συνοδεύουν τις ορθογραφικές μεταγραφές και τη μορφολογικές αναλύσεις.

[Τεχνικά] Μια αξιολόγηση της βάσης από ερευνητές που έχουν ήδη δουλέψει με τη βάση, καθώς και από προσωπικό που έχει εμπειρία από άλλες ηλεκτρονικές βάσεις

[Γλωσσολογικά] Έναρξη διερεύνησης του σώματος όλων των διαλεκτικών δεδομένων με τη χρήση του μορφολογικού αναλυτή, για παράδειγμα με το TOOLBOX, ώστε να δημιουργήσουμε ένα καλό λεξικό.

[Γλωσσολογικά] Εμπλουτισμός του διαλεκτικού υλικού, τόσο προφορικού, όσο και γραπτού, με την οργάνωση νέων αποστολών και συλλογών υλικού, καθώς και την ψηφιοποίηση του γραπτού υλικού που έχουμε στην κατοχή μας.

[Έρευνα] Σχεδιάζουμε την έκδοση λεξικών, λεξιλογίων και γραμματικών για τις διαλέκτους που έχουμε μεγάλο εύρος προφορικού υλικού.

[Έρευνα] Επιπλέον χορήγηση ερευνητικών προσπαθειών για οικονομική υποστήριξη με στόχο τη βελτίωση και εξέλιξη της ηλεκτρονικής βάσης GREE.D.

[Έρευνα] Χρήση της βάσης δεδομένων ως βοηθητικό εργαλείο για τη μελλοντική διαλεκτική έρευνα για διάφορα φωνολογικά και μορφολογικά φαινόμενα, τα οποία εντοπίζονται δια-διαλεκτικά και αποτελούν σημαντικότερο αρωγό για την παραγωγή άρθρων και μονογραφιών για τις διάφορες νεοελληνικές διαλέκτους.

[Έρευνα] Επικοινωνία και συνεργασία με τη διεθνή γλωσσολογική κοινότητα, ώστε να παρέχουμε τη δυνατότητα πρόσβασης σε ελληνικά διαλεκτικά δεδομένα και παράλληλα να διατηρήσουμε και να διασώσουμε μια εξαιρετικά σημαντικά πολιτιστική κληρονομιά.

Βιβλιογραφία

Allen, W., Beal, J., Corrigan, K., Maguire, Moisl, H. (2007). *A Linguistic 'Time Capsule': The Newcastle Electronic Corpus of Tyneside English*. Creating and digitalizing Language Corpora Vol.2 (edited by Beal J. et al.). Palgrave MacMillan Publication, pp. 16-48.

- Anderson, J., Beavan, D., Kay, C. (2007). *SCOTS: Scottish Corpus of Texts and Speech*, Creating and digitalizing Language Corpora Vol.1 (edited by Beal J. et al.). Palgrave MacMillan Publication, pp. 17-34.
- Anderwald, L., Wagner, S. (1997). *FRED – The Freiburg English Dialect Corpus: Applying Corpus-Linguistic Research Tools to the Analysis of Dialect Data*. Creating and digitalizing Language Corpora Vol.1 (edited by Beal J. et al.). Palgrave MacMillan Publication.
- Barbiers, S., Cornips, L., Kunst, J.-P. (2007). *The Syntactic Atlas of the Dutch Dialects (SAND): A Corpus of Elicited Speech as an On-line Dynamic Atlas*. Creating and digitalizing Language Corpora Vol.1 (edited by Beal J. et al.). Palgrave MacMillan Publication.
- Dipper, S., Goetze, M., Skopeteas, S. (2007). *Information Structure in Cross-linguistic corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure*. ISIS Working papers of the SFB 632.
- Gordon, E., Maclagan, M., Hay, J. (2007). *The ONZE Corpus*. Creating and digitalizing Language Corpora Vol.2 (edited by Beal J. et al.). Palgrave MacMillan Publication, pp. 82-104.
- MacWhinney, B. (2007). *The Talkbank Project*. Creating and digitalizing Language Corpora Vol.1 (edited by Beal J. et al.). Palgrave MacMillan Publication.
- Ralli, A. (2006). Syntactic and Morphosyntactic Phenomena in Modern Greek Dialects: The State of the Art. 2007. *Journal of Greek Linguistics 2006*: 121-159. Ακαδημία Αθηνών (1933-). *Ιστορικό Λεξικό της Νέας Ελληνικής Γλώσσας, της τε Κοινώς Ομιλουμένως και των Ιδιωμάτων*. Αθήνα
- Ιστορικό Αρχείο Ελλήνων Προσφύγων Καλαμαριάς Θεσσαλονίκης, (http://www.kalamaria.gr/index.php?option=com_content&task=view&id=85&Itemid=599)
- Κοντοσόπουλος, Ν. (2006)⁴. *Διάλεκτοι και ιδιώματα της Νέας Ελληνικής*. Αθήνα: εκδ. Γρηγόρης.
- Μηνάς, Κ. (2003), *Η γλώσσα των Δημοσιευμένων Μεσαιωνικών ελληνικών εγγράφων της Κάτω Ιταλίας και της Σικελίας*. (επανέκδοση από Ι.Α.Ν.Ε.), Αθήνα.
- Καραναστάσης, Α. (1986-1992). *Ιστορικό Λεξικό των Ελληνικών Ιδιωμάτων της Κάτω Ιταλίας*. Τόμοι Α-Ε. Αθήνα.
- Καραναστάσης, Α. (1992). *Γραμματική των Ελληνικών Ιδιωμάτων της Κάτω Ιταλίας*. Αθήνα
- Κωστάκης, Θ. (1986-1987). *Λεξικό της Τσακωνικής Διαλέκτου*. Τόμοι Α-Γ, Αθήνα.
- Ράλλη, Α. (to appear). *Λεξικό των Ιδιωμάτων Κυδωνιών, Μοσχονησίων και Ανατολικής Λέσβου*. Πανεπιστήμιο Πατρών: Εργαστήριο Νεοελληνικών Διαλέκτων.
- Ίδρυμα Μανώλη Τριανταφυλλίδη (<http://ins.web.auth.gr/english.htm>)

Η έρευνα για τις νεοελληνικές διαλέκτους στο Ινστιτούτο Νεοελληνικών Σπουδών

ΓΙΩΡΓΟΣ ΠΑΠΑΝΑΣΤΑΣΙΟΥ
Ινστιτούτο Νεοελληνικών Σπουδών,
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
papana@lit.auth.gr

Θα ήθελα κατ' αρχάς να ευχαριστήσω τους διοργανωτές του συνεδρίου, την κ. Αγγελική Ράλλη, τον κ. Mark Janse και τον κ. Brian Joseph, για την πρόσκληση που μου απηύθυναν να συμμετάσχω στη σημερινή συνάντηση, εκπροσωπώντας το Ινστιτούτο Νεοελληνικών Σπουδών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, και να παρουσιάσω με τον τρόπο αυτό τις ερευνητικές του δραστηριότητες που σχετίζονται με τις νεοελληνικές διαλέκτους.

Είναι γνωστό στο κοινό του Διεθνούς Συνεδρίου Νεοελληνικών Διαλέκτων και Γλωσσολογικής Θεωρίας ότι εδώ και μια επταετία το Ινστιτούτο Νεοελληνικών Σπουδών, ύστερα από πρόταση του Χρ. Τζιτζιλί, ανέλαβε την πρωτοβουλία να ετοιμάσει, υπό την εποπτεία του ίδιου, ένα συλλογικό έργο με τίτλο *Νεοελληνικές διάλεκτοι*, το οποίο φιλοδοξεί να καλύψει ένα υπαρκτό κενό στον τομέα της μελέτης τους.

Η πρωτοβουλία αυτή του Ινστιτούτου ξεκίνησε ύστερα από τη σχεδόν αυτονόητη διαπίστωση ότι δεν υπάρχει μια σύγχρονη, γενική, εκτενής και συστηματική περιγραφή όλων των νεοελληνικών διαλέκτων και ιδιωμάτων. Η *Νεοελληνική γραμματική, Ιστορική εισαγωγή* του Μ. Τριανταφυλλίδη (1939) και οι *Διάλεκτοι και ιδιώματα της νέας ελληνικής* του Ν. Κοντοσόπουλου (2001), έργα χρησιμότετα και τα δύο, αποτελούν τις μοναδικές ως τώρα απόπειρες περιγραφής όλων των νεοελληνικών διαλεκτικών ποικιλιών, δίνοντας πολύτιμες πληροφορίες όσον αφορά την ταξινόμηση και τα βασικά χαρακτηριστικά τους, ο στόχος τους όμως, ο οποίος αντανακλάται και στην έκτασή τους, απέχει πολύ από τη συστηματική και ολοκληρωμένη περιγραφή κάθε διαλέκτου ή διαλεκτικής ομάδας σε όλα τα επίπεδα της γλωσσολογικής ανάλυσης. Διαφορετικός είναι ο στόχος του επίσης σημαντικού, πρόσφατου συλλογικού τόμου *Διαλεκτικοί θύλακοι της ελληνικής γλώσσας*, σε επιμέλεια του Τάσου Χριστίδη, που κυκλοφόρησε το 1999 από το Κέντρο Ελληνικής Γλώσσας.

Στην απόφαση του Ινστιτούτου για την προετοιμασία του έργου έπαιξε, επίσης, ρόλο το γεγονός ότι οι νεοελληνικές διάλεκτοι δεν έτυχαν όλες της ίδιας προσοχής, με αποτέλεσμα να υπάρχουν σημαντικές διαφοροποιήσεις ως προς την πληρότητα με την οποία έχουν μελετηθεί. Ορισμένες ευτύχησαν να περιγραφούν με επάρκεια και από σπουδαίους μελετητές, είτε συνολικά είτε τουλάχιστον σε βασικά επίπεδα της γλωσσολογικής ανάλυσης, κυρίως στη φωνητική και το λεξιλόγιο. Τέθηκαν έτσι από παλαιά στο επίκεντρο του ενδιαφέροντος και λειτούργησαν έκτοτε ως όχημα για την ανάπτυξη της νεοελληνικής διαλεκτολογίας. Άλλες μελετήθηκαν πιο αποσπασματικά, ενώ ορισμένες είτε έγιναν ευρύτερα γνωστές σχετικά πρόσφατα είτε για διάφορους λόγους παραμελήθηκαν. Συνδυασμοί όλων αυτών των παραγόντων οδήγησαν ώστε οι μελέτες, και συνακόλουθα οι γνώσεις, που έχουμε σήμερα π.χ. για την ποντιακή να είναι δυσανάλογες σε σχέση με αυτές που έχουμε για άλλες διαλέκτους. Η ισότιμη αντιμετώπιση όλων ήταν, επομένως, ένα ζητούμενο.

Τέλος, δεν θα πω κάτι καινούργιο, ισχυριζόμενος ότι και σε αυτό το πεδίο της γλωσσολογικής έρευνας διαθέτουμε σήμερα, για όλες τις ευρωπαϊκές γλώσσες, αναλυτικές περιγραφές των διαλέκτων τους, που ξεκινούν από παρόμοιες αρχές με αυτές που τέθηκαν στο υπό προετοιμασία έργο. Όλες οι παρατηρήσεις αυτές παραπέμπουν στο γεγονός ότι το έργο αυτό είναι ένα επιστημονικό ζητούμενο στον ευρύτερο χώρο της ελληνικής γλωσσολογίας.

Η συνειδητοποίηση αυτών των ελλείψεων έπαιξε ρόλο στις αποφάσεις σχετικά με τον χαρακτήρα, τη μορφή και τη δομή που το έργο θα έπρεπε να πάρει. Στόχος του είναι αφενός να συγκεντρώσει και να παρουσιάσει τα πορίσματα της έως τώρα νεοελληνικής διαλεκτολογικής έρευνας, αποτελώντας μια επισκόπησή της, αφετέρου όμως να αποτελέσει την αφετηρία για καινούργιους επιστημονικούς προβληματισμούς, με βάση νεότερα δεδομένα και λαμβάνοντας υπόψη, επίσης, τη σύγχρονη γλωσσολογική θεωρία. Και το σημαντικό – αλλά ταυτόχρονα και το δύσκολο κατά τη σύνταξη και την προετοιμασία του – είναι ότι, εκτός από το ότι λαμβάνει υπόψη και υλικό που συλλέχθηκε πρόσφατα – αποτελεί σε μεγάλο βαθμό αποτέλεσμα σύγχρονης, και συχνά διαφορετικής, ανάλυσης των ήδη γνωστών διαλεκτικών δεδομένων. Το επιχειρούμενο αποτέλεσμα αποτελεί τελικά συνδυασμό της συγχρονικής ανάλυσης και της διαχρονικής ερμηνείας, έναν συνδυασμό που – ας μου επιτραπεί να το πω – όχι μόνο δεν είναι πάντοτε αυτονόητος αλλά συχνά πολεμήθηκε υπέρ μιας καθαρά είτε συγχρονικής είτε διαχρονικής ανάλυσης.

Η διάρθρωση του έργου προκύπτει, σε μεγάλο βαθμό, από αυτούς τους στόχους. Προβλέπεται δίτομο, με τον πρώτο τόμο να αποτελεί μια εκτενή Εισαγωγή και τον δεύτερο να εξετάζει αναλυτικά τις επιμέρους διαλέκτους και τα ιδιώματα. Ο δεύτερος τόμος θα κυκλοφορήσει πρώτος, βρίσκεται πλέον στην τελική ευθεία για την ολοκλήρωσή του και σε αυτόν συμμετέχουν δεκαέξι Έλληνες και ξένοι συγγραφείς. Περιλαμβάνει μονογραφίες έκτασης περίπου 80 σελίδων η καθεμιά για την ποντιακή, την κραιοαζοφική, την καππαδοκική, τα υπόλοιπα μικρασιατικά ιδιώματα, την κυπριακή, την τσακωνική, την πελοποννησιακή, τημανιάτικη, τημεγαροκουμιώτικη, την κρητική, την επτανησιακή, τη δωδεκανησιακή, την κυκλαδίτικη, την κατωιταλική και τα βόρεια ιδιώματα, ενώ υπάρχουν και ξεχωριστές μικρότερης έκτασης συμβολές, π.χ. για τα ιδιώματα της Χίου.

Κάθε συμβολή ξεκινάει με την περιγραφή του γεωγραφικού και του ιστορικού πλαισίου εντός των οποίων αναπτύχθηκε κάθε διάλεκτος, και στη συνέχεια παρέχει μια επισκόπησή της ως τώρα σχετικής έρευνας. Όσον αφορά τα γενικά χαρακτηριστικά κάθε διαλέκτου, επιλέγονται 25 βασικά φωνητικά, μορφολογικά, συντακτικά και λεξιλογικά ισόγλωσσα, με τα οποία επιχειρείται η ταξινόμηση και η τοποθέτηση των ιδιωμάτων στον διαλεκτικό χώρο. Ακολουθούν ειδικότερα χαρακτηριστικά κάθε διαλέκτου, αναφορές στις παλαιότερες γνωστές μορφές της και στη συνέχεια δίνεται μια εικόνα της – πάντοτε υπαρκτής – ενδοδιαλεκτικής διαφοροποίησης.

Η συστηματική περιγραφή που ακολουθεί ξεκινάει από τη φωνητική και τη φωνολογία, περνάει στη μορφολογία, στη σύνταξη, στο λεξιλόγιο, στη φρασεολογία, στην παραγωγή και στη σύνθεση, στις σχέσεις με άλλες διαλέκτους, και ολοκληρώνεται με την επισκόπησή της σημερινής κατάστασης του ιδιώματος. Ύστερα από μια σύντομη αναφορά στην πιθανή σύγχρονη παραγωγή γραπτού λόγου σε κάθε ιδίωμα, κάθε συμβολή ολοκληρώνεται με την παράθεση σχολιασμένων διαλεκτικών κειμένων και, βέβαια, της σχετικής βιβλιογραφίας.

Ο δεύτερος αυτός τόμος, που προβλέπεται να έχει έκταση 1.200 σελίδων, βρίσκεται στο τελευταίο στάδιο της προετοιμασίας του, καθώς ολοκληρώνεται η φιλολογική του επιμέλεια, ενώ στο αρχικό στάδιο βρίσκεται η κατάρτιση των ευρετηρίων λέξεων και όρων. Απρόβλεπτες καθυστερήσεις, εν μέρει δικαιολογημένες με βάση την έκταση του εγχειρήματος, είχαν ως συνέπεια να υπερβούμε τις αρχικές προθεσμίες που είχαμε θέσει για την έκδοσή του, είμαστε όμως αισιόδοξοι για την ολοκλήρωση του εγχειρήματος και για την έκδοση του τόμου την επόμενη χρονιά. Ο πρώτος τόμος, αυτός της Εισαγωγής, αναμένεται να κυκλοφορήσει αργότερα. Θα περιέχει εκτενή ανάλυση των φαινομένων που παρουσιάζονται στον ελληνόφωνο διαλεκτικό χώρο, αφενός ιστορική και αφετέρου τυπολογική, και θα τον υπογράψει ο Χρ. Τζίτζιλής. Ειδικότερα, τα 25 χαρακτηριστικά βάσει των οποίων γίνεται η ταξινόμηση των νεοελληνικών διαλέκτων – και όχι μόνο αυτά – θα αναλύονται διεξοδικά και θα ερμηνεύονται ιστορικά.

Τα πράγματα όμως δεν σταματούν εδώ. Και βέβαια δεν θα σταματήσουν με την έκδοση του έργου. Ήδη αποφασίστηκε η έκδοση από το Ινστιτούτο του *Γλωσσικού*

άτλαντα της Δωδεκανήσου, και είναι η πρώτη φορά που ανακοινώνεται αυτό επίσημα, τον οποίο υπογράφει ο Κ. Μηνάς. Το ζητούμενο όμως δεν είναι μόνο οι εκδόσεις. Και αυτό γιατί η βασικότερη πτυχή κάθε επιστημονικής δραστηριότητας είναι η έρευνα. Εν μέρει λοιπόν βάσει του αρχικού σχεδιασμού, εν μέρει όμως και λόγω της συνεχούς αναθεώρησης των στόχων, και βέβαια σε συνδυασμό με τις ανάγκες ελέγχου, διόρθωσης και ομοιόμορφης παρουσίασης των ίδιων των συμβολών, και σύνταξης της Εισαγωγής, το Ινστιτούτο αισθάνθηκε ότι θα έπρεπε και θα μπορούσε να συμβάλει σε μια μονιμότερη ανάπτυξη των διαλεκτολογικών σπουδών και γενικά του επιστημονικού ενδιαφέροντος για τις νεοελληνικές διαλέκτους.

Ξεκινώντας σχεδόν από το μηδέν, συγκροτήθηκε σιγά σιγά μια αρκετά καλά ενημερωμένη διαλεκτολογική βιβλιοθήκη, η οποία αγγίζει τους 600 τίτλους και εμπλουτίζεται συνεχώς. Δημιουργήθηκε, επίσης, αρχείο άρθρων και ανατύπων, που σήμερα αριθμεί περίπου 3.200 καταχωρήσεις (σε ένα σύνολο 6.200, που αφορούν την ιστορία της ελληνικής). Και βέβαια συγκροτήθηκε ένα αρχείο με προφορικές καταγραφές από διάφορα μέρη του ελληνόφωνου χώρου, με έμφαση σε εκείνα από τα οποία η παλαιότερη έρευνα δεν είχε καταφέρει να συλλέξει αρκετά στοιχεία. Στο πλαίσιο αυτό οργανώθηκαν αποστολές σε διάφορες περιοχές, από την Τσακωνιά ως τη Βουλγαρία, για να αποκτηθεί νέο πρωτογενές υλικό, με αποτέλεσμα αυτή τη στιγμή στο Ίδρυμα να στεγάζονται 350 περίπου κασέτες. Το φωνητικό αυτό υλικό προς το παρόν δεν είναι προσπελάσιμο από άλλους ερευνητές, θα εξεταστεί όμως αργότερα η δυνατότητα περαιτέρω αξιοποίησής του.

Από την άλλη, ένας δεύτερος αλλά εξίσου σημαντικός – αν όχι σημαντικότερος – στόχος ήταν και είναι να δημιουργηθεί ένας μικρός πυρήνας ερευνητών που θα ασκηθούν στη μελέτη των διαλέκτων, θα αποκτήσουν τα εφόδια για τη μελέτη τους και τον προβληματισμό για τους σχετικούς τρόπους έρευνας. Ο πυρήνας αυτός άρχισε να συγκροτείται πριν από 7-8 χρόνια, ως τώρα έχουν περάσει από την εκπαίδευση αυτή 10 ερευνητές και μεταπτυχιακοί φοιτητές και σήμερα αριθμεί 6 μέλη που εργάζονται σε καθημερινή βάση. Το Ινστιτούτο προσπαθεί από αυτή την άποψη να εκπαιδεύσει νέους ανθρώπους και να δημιουργήσει μια ομάδα διαλεκτολόγων, οι οποίοι θα συνδυάζουν τις μεταπτυχιακές ή διδακτορικές σπουδές τους με μια χρήσιμη και εποικοδομητική εργασία.

Συνεχίζοντας, κυρίες και κύριοι, την παράδοση του Μανόλη Τριανταφυλλίδη, στον οποίο οι διαλεκτικές σπουδές οφείλουν πολλά, και όχι μόνο λόγω της *Ιστορικής εισαγωγής στη Νεοελληνική γραμματική της δημοτικής*, το Ινστιτούτο αποκτά, τόσο με την επικείμενη έκδοση του έργου με τον τίτλο *Νεοελληνικές διάλεκτοι* όσο και με τη δημιουργία υποδομών για τη συνέχιση της μελέτης τους, μια θέση ανάμεσα στα ελάχιστα ερευνητικά κέντρα που ασχολούνται με το ζήτημα αυτό και φιλοδοξεί αυτή τη θέση να τη βελτιώσει στο μέλλον.

Research Center for Modern Greek Dialects - Historical Dictionary

CHRISTINA BASSEA-BEZANTAKOU
Academy of Athens

0. Introduction

This paper aims to give an overall presentation of the Research Center for Modern Greek dialects - Historical Dictionary of the Academy of Athens (henceforth RCMGD), by providing a short description of its research, lexicographical and publication activities. Detailed information is available at the Center's website, <http://www.academyofathens.gr/ilne/>. Special emphasis is accorded to the presentation of the content, the methodology and the problems of compilation of the Historical Dictionary of Modern Greek (*Ιστορικὸν Λεξικὸν τῆς Νέας Ἑλληνικῆς Γλώσσης τῆς τε κοινῶς ὁμιλουμένης καὶ τῶν ἰδιωμάτων*), which constitutes the RCMGD's main publication activity.

1. Foundation- History

The RCMGD was established in 1908 on the initiative of Georgios Chatzidakis, the founder of the linguistic science in Greece⁴, and is one of the first research centers to be placed under the auspices and superintendence of the Academy of Athens (1927).

Its first appellation was "Research Center for the Compilation of the *Historical Dictionary of Modern Greek, both of the common language and its dialects*", which described exactly the main purpose of the Centre. That is, its initial purpose was the compilation of a unified dictionary of the Greek language, from antiquity until modern times, which would be called "historical" because it would provide information on the history of words, i.e. their phonological, morphological and semantic evolution along the time axis. It would constitute "proof of the linguistic unity of the nation through the centuries, and a monument to the immortality of the Greek race" ("μνημεῖον τῆς γλωσσικῆς ἐνότητος τοῦ ἔθνους διὰ πάντων τῶν αἰώνων καὶ μνημεῖον τῆς ἀθανασίας τῆς ἑλληνικῆς φυλῆς")⁵. This over-ambitious initial purpose, understandable within the framework of late 19th c. nationalism and romantism (see Giakoumaki, Karantzi & Manolessou 2004), was of course unachievable with the means available at the time. Thus, the scope of the dictionary was soon revised and reduced to the spoken Modern Greek language, "both the commonly spoken language and its dialects".⁶

The slow rate of progress in the creation of the *Historical Dictionary*, as well as the new standards required by modern linguistics and dialectology, to which it must conform, led to yet another revision of its purposes, and in 2003 the Center was re-named "Research Center for Modern Greek Dialects- Historical Dictionary".

2. Archives of the RCMGD

Given that the surviving linguistic material of the ancient, medieval and early modern period exists in written form, the first step for the compilation of the *Historical*

⁴ For details on this major figure in the history of Greek linguistics see Vayakakos (1977).

⁵ In the "Prolegomena" of the first volume of the dictionary (*HD*, vol. 1, p. η') the type and the content of the dictionary under preparation is described as follows: "all ancient linguistic items, i.e. all words, all word-forms and sounds, all meanings etc. must be followed down to their final disappearance, if they have been lost, or down to the present, if they have survived, and conversely, all new linguistic items, be they words, word-forms, sounds, meanings, phrases etc. must be followed back to their first appearance in history"

⁶ See also the 1925 Introduction to the Liddell-Scott *Greek-English Lexicon* (LSJ: vi) where the initial scope of the Historical Dictionary of Modern Greek and its subsequent revision are discussed.

Dictionary of Modern Greek was the coverage of the spoken modern language, i.e. the local spoken varieties in all Greek-speaking areas, inside and outside Greece. Therefore, the creation of the Historical Dictionary archives begins with the collection of modern Greek vernacular spoken language, on the one hand from printed sources (dictionaries, glossaries, scholarly and literary periodicals and books containing samples of spoken language) and on the other from manuscript collections of linguistic material compiled by the Linguistic Society of Athens (εν Αθήναις Γλωσσική Εταιρεία), the Greek Philological Society of Constantinople (Ελληνικός Φιλολογικός Σύλλογος Κωνσταντινουπόλεως) or coming from the private archive of Michael Deffner (1848-1934), one of the first researchers of the Tsakonian dialect. In parallel, regular fieldwork collection of linguistic material from various Greek-speaking areas was conducted (and is still being conducted) by the researchers working for the Historical Dictionary. As a result, the RCMGD archives nowadays contain the following material:

2.1. Manuscript collections archive:

There are 1680 manuscript collections (transcriptions) of spoken language in the Historical Dictionary archive, whose provenance is the following:

- a) data collected/ transcribed during fieldwork by the researchers of the Center from 1908 until today.
- b) data submitted to the annual linguistic competitions of the Linguistic Society of Athens and, nowadays, to the linguistic competition organized by the Academy of Athens on behalf of the RCMGD
- c) 191 manuscripts donated by the Greek Philological Society of Constantinople to the RCMGD, compiled around the middle of the 19th c. (the oldest is dated 1854, but the linguistic situation these mss. reflect is much older). Most of these manuscripts are in a very poor condition of preservation and are kept in a separate sub-archive.
- d) linguistic data collected and donated by private individuals.

The archive material of the RCMGD is of very considerable value, since it constitutes the oldest and by far largest collection of data on the Modern Greek dialectal varieties. On the contrary, the representation of common (standard) Modern Greek in the the RCMGD archives is much more restricted. The content of most data collections in the archives, especially the older ones, shows that the interest of the collector usually lay in the recording of words which were unfamiliar to the speaker of the standard language; and the further excerpting and use of these collections as sources of lexicographical material for the compilation of the Historical Dictionary was to a certain extent based on the same principle.

2.2. Card slip archive

The card-slip archive consists of more than 4.000.000 card-slips indexed by lemma in alphabetical order (for example, the lemma *γριά* 'old woman' includes the dialectal forms *γραιία, εγραιία, γραιίε, γρία, γρία, γρ^αιά, γιργιά, γριτζιά, ρκά, γρέ...* as well as the plural forms *γριές, γρες, γραιάδοι* etc.)⁷. It has been excerpted from:

- a) the manuscript data collections described above
- b) printed sources containing vernacular and especially dialectal material.

The card-slip archive is divided in three parts: i) the excerpts used for the compilation of the already printed volumes of the *HD* (α-δαχτυλωτός) ii) the appendix containing additional material for the printed lemmata, excerpted after the publication of the first volumes and iii) the main body of the card-slip archive, containing material from the lemma *δε* and following (up to *ωωχά*). Sections (ii) and (iii) are constantly being added to. The card-slips are hand-written, and most of the older ones are especially problematic due to the unsystematic way of recording of the data (abbreviation of the source, phonetic

⁷ Concerning the methodology of choosing and compiling a lemma in a historical dictionary, see Bassea-Bezantakou (1997, 2006).

transcription) and sometimes to the bad state of preservation of the paper, ink etc. To give a characteristic example, many of the older slips, especially those created during the period of the German occupation of Greece (1941-45) are written on the margin of paper-slips coming from book pages, newspapers, voting papers etc.

2.3. Donated archives.

Three archival bodies have been donated in their integrity to the RCMGD:

- a) the archive of M. Deffner (Tsakonian)
- b) the archive of St. Karatzas (Euboea and other areas)
- c) the archive of A. Karanastasis (S. Italy)

2.4. Sound archive

This part of the RCMGD archives is currently in the process of being created, through the digitization of older sound recordings preserved in various mediums, such as magnetic reels, cassette tapes, videotapes, and vinyl records. These sound files contain live recordings of oral dialectal material, of various local provenance, collected from 1930 onwards. The digitization is being carried out with the valuable assistance of the Laboratory of Modern Greek Dialects of the University of Patras.

2.5. Toponyms archive

Since 1984, an electronic database of Greek place-names is being constructed and constantly enriched. The data is excerpted from the manuscripts archive. For more information, see Afroudakis (2001).

3. The Historical Dictionary of Modern Greek

3.1 Aims-Scope-Problems

As already discussed, the compilation of a dictionary for the contemporary spoken language (both the common language and its dialects), which would also provide the historical overview of each lexical item, was the ultimate aim set down in 1910 by G. Chatzidakis, the dictionary's founder. The "contemporary spoken language" was considered to be the linguistic form in use from 1800 onwards, and in particular "the language of the people, and not the written demotic used by many" (τῆς γλώσσης τοῦ λαοῦ καὶ οὐχὶ τῆς ὑπὸ πολλῶν γραφομένης μάλιστα σήμερον δημοτικῆς), with the addition of lexical items "preserved in lexicographical or literary works, so long as they are of genuine vernacular form" and with the exception of such words "if they appear to be nonce formations" ("πᾶσαι αἱ λέξεις, αἱ ὁποῖα παραδίδονται ὑπὸ λεξικογράφων ἢ λογοτεχνῶν, ἀρκεῖ νὰ φέρουν γνησίαν δημῶδη μορφήν... ἀποκλείονται δὲ λέξεις λεξικογράφων ἢ λογοτεχνῶν, αἱ ὁποῖα φαίνονται πρόχειρα κατασκευάσματα" (*HD*, Prolegomena to vol. 1, p. ιζ'). The contemporary spoken language was thus defined on the basis of a single criterion: that the words in question be not ignored by "the people", a criterion of doubtful accuracy and practical applicability which needs no further comments. Irrespective of how rich the archival material is, it is impossible for it to be analysed synchronically, considering that it was collected in different periods, following different criteria, and especially since it was collected over such a long period (from 1916 and still ongoing), from informants answering to variable presuppositions as to educational level, age, gender, social class etc. In addition, the temporal distance between the initial conception of the *Historical Dictionary* and the present day has inevitably brought about a discrepancy between its original principles and purposes and the current status and methodology of research of the Modern Greek dialects and the Modern Greek linguistic reality.

The specification of the scientific research domain of the Center as being the Modern Greek dialects, as expressed by its renaming in 2003 to "Research Center for Modern Greek Dialects- Historical Dictionary" (Κέντρον Ερεῦνης Νεοελληνικῶν Διαλέκτων καὶ

Ιδιωμάτων -IANE), gives a more accurate picture of the lexicographical work being carried out by its researchers. That is, the lexicographical purpose of the Historical Dictionary is first of all the charting of the history of the Greek language through the investigation of its dialects, since true linguistic history is most often detectable through dialectal material and obscured in the standard language. Through spatial linguistic variation it is possible to establish the changes that affect the later Greek language, not only on the level of lexicon and semantics, but on the phonological and morphological level as well. It is thus the very nature of the material that requires a double form of investigation (historical and comparative-dialectal) for the compilation of the Historical Dictionary. This double attention to both language history and dialectal variation is a unique, although necessary, practice in the domain of international lexicography, since usually the focus of major national lexicographical projects is either historical or dialectological.⁸

To conclude, the aim of the Historical Dictionary is not to hoard the whole thesaurus of the Modern Greek language from 1800 onwards, but to ensure the preservation of the dialectal fund of the Greek language (not only as lexical items but as meanings as well) and to investigate the history of the Modern Greek language. The complete coverage of all aspects of the lexicon of Standard Modern Greek is the domain of other well-known lexicographical projects, such as the already published dictionaries LNE and LKN, as well as the new Dictionary of Current Greek which is will shortly appear under the auspices of the Academy of Athens (see Charalambakis 2009 for details).

3.2. Sources of Historical Dictionary

The provenance of the material used to compile the HD is the following: 1) primary/oral sources, which provide mainly dialectal material (see above under archives of the RCMGD) and 2) secondary/written sources, which provide material both on standard Modern Greek and its dialects. The secondary sources of dialectal material can be divided into:

- a) Linguistic sources: local dictionaries and glossaries, linguistic treatises and papers concerning a certain dialect, phenomenon or lexical item
- b) Folklore sources: collections of folk-tales, songs, proverbs, customs
- c) Literary sources: literary works written in a certain dialect

Correspondingly, the secondary sources for standard Modern Greek can be divided into:

- a) Linguistic sources: the major and minor dictionaries of Modern Greek, the available electronic corpora of Modern Greek (Hellenic National Corpus and the Corpus of Greek Texts (ΕΘΕΓ and ΣΕΚ)⁹, linguistic treatises and papers concerning a certain phenomenon or lexical item
- b) Literary sources: literary works written in standard Modern Greek, electronic corpora containing literary works of Modern Greek

Additionally, in order to document the history of the Modern Greek vocabulary more fully, the HD consistently uses written sources containing material from earlier periods of Greek such as Lexica and Grammars of Ancient, Koine and Medieval Greek, electronic corpora of Ancient, Koine and Medieval Greek (mainly the TLG, www.tlg.uci.edu), etc.

4. New principles in the research and lexicographical activities of the RCMGD

Within a general framework of modernization in the processes of compilation of the HD according to the principles and presuppositions of modern Lexicography and Dialectology, the following innovations are currently under way:

⁸ For a comparison of the Historical Dictionary of Greek with similar international lexicographical projects see Giakoumaki & Karantzi & Manolessou (2004).

⁹ See their web-pages, <http://hnc.ilsp.gr/> and <http://sek.edu.gr/> respectively.

a) phonetic transcription of all linguistic forms with the International Phonetic Alphabet and use, where necessary, of a set of special phonetic symbols for the transcription of dialectal texts in the Greek alphabet. These symbols are included in a font designed and created specifically for this purpose, the *Athens Academy Greek Fonts*, and a table of one to one correspondence with the phonetic symbols of the IPA allows their appropriate use.

b) The metalanguage of the HD will no longer be the linguistic form of the previous volumes, i.e. the katharevousa (or occasionally an even more archaic form) but Standard Modern Greek, obeying determined grammatical and orthographical principles.

c) use of electronic databases: for the facilitation and the speeding-up of the compilation process, a number of digital/electronic databases have been constructed, namely:

i) a database containing detailed information on the manuscript collection, allowing complex searches by author, date of collection, area of provenance, or content of the manuscript. Especially for material collected after 1922 from Asia Minor refugees, special attention is paid to the recording both of the area of provenance and the area of relocation. The contribution of this database is crucial, since it will ensure the reliability of the archival material and will allow the user of the HD to conceive the material both in its synchronic and its diachronic dimension. This information will be provided also in the printed form of the HD.

ii) a database containing the place-names (and their abbreviations) used for the description of the provenance of linguistic material in the HD (see Bassea-Bezantakou 2001). The place-names are fully defined geographically according to prefecture, county, older appellation etc., on the basis of the official Administrative Division of Greece of 1996.¹⁰

iii) Digitised archive of scanned mss and digitized archive of sound recordings (under construction)

iv) Electronic edition of the first published volumes of the *HD* (α-δαχτυλωτός) (under construction)

v) Bibliographical database containing a) updated and cross-checked bibliography of the published volumes and b) catalogue of the research library of the RCMGD, which is the richest one in Greece in the domain of dialectology

vi) Digitised archive of scanned printed secondary sources (dictionaries, dialectal glossaries and treatises)

vii) Electronic database containing in summary form the contents of the card-slip archive and allowing complex searches by lemma or linguistic form. This database, although not providing a full electronic and accurate transcription of the archive, is an invaluable first step in the detection and location of dialectal linguistic forms.

5. Publications of the RCMGD

Apart from the Historical Dictionary of Modern Greek (in its new form) and its other previous publications, the RCMGD is also responsible for the following publications:

- the RCMGD's journal *Λεξικογραφικόν Δελτίον* (the 26th volume is currently in preparation)

- the collective volume series *Νεοελληνική Διαλεκτολογία* (the 6th volume is currently in preparation)

- the new dictionary of dialectal archaisms, by the ex-director of the RCMGD, Dr D. Krekoukias, which will shortly appear in 2 volumes (*Αρχαϊσμοί στα Νεοελληνικά Ιδιώματα*)

¹⁰ *Γεωγραφικός Κώδικας της Ελλάδος* Υπουργείο Εσωτερικών, Δημόσιας Διοίκησης και Αποκέντρωσης, Athens 1997.

6. Conclusion

It is hoped that this short overview of the research and lexicographical work being carried out at the RCMGD has allowed the reader to form an optimistic view of the new perspectives which the of the research and lexicographical work being carried out at the RCMGD allows, not only concerning the publication of the Historical Dictionary but in general concerning the scientific progress in the study of both the Modern Greek dialects and the history of the Greek language.

References

- Afroudakis, A. (2001): "Εργαλεία' που αναπτύχθηκαν από το «Αρχείο τοπωνυμίων και κυρίων ονομάτων» στο πλαίσιο του Προγράμματος ΑΜΗΤΟΣ". *Λεξικογραφικόν Δελτίον* 23 (2001): 159-225.
- Bassea-Bezantakou, Ch. (1997): "Σπουδή συντάξεως ενός λήμματος του Ιστορικού Λεξικού", In *Φιλερήμου Αγάπησις, Τιμητικός Τόμος για τον Καθηγητή Αγ. Τσοπανάκη*, Rhodes: Stegi Grammaton kai Technon, 419 - 428.
- Bassea-Bezantakou, Ch. (2001): "Μηχανοργάνωση του πίνακα τοπωνυμίων του ΙΑΝΕ στο πλαίσιο του Προγράμματος ΑΜΗΤΟΣ". *Λεξικογραφικόν Δελτίον* 23 (2001): 109-129.
- Bassea-Bezantakou, Ch. (2006): "Dialect Dictionaries: synchronic and diachronic dimensions of the material", In *Filia. Studies in Honour of Bo-Lennart Eklund*. Lund: Media Tryck, 27-40.
- Charalambakis, Ch. (2009): Το νέο Χρηστικό Λεξικό της Νεοελληνικής της Ακαδημίας Αθηνών. In *Proceedings of the 8th International Conference on Greek Linguistics, Ioannina, August 30th-September 2nd 2007*, 1263-1282. [e-book]
- Giakoumaki, E., Ch. Karantzi & I. Manolesou (2004): ΙΑΝΕ και ηλεκτρονική λεξικογραφία. *Proceedings of the 6th International Conference on Greek Linguistics, Rethymnon 18-21 September 2003*, [e-book]
- HD = Historical Dictionary (1933-) *Ιστορικόν Λεξικόν τῆς Νέας Ἑλληνικῆς Γλώσσης τῆς τε κοινῶς ὁμιλουμένης καὶ τῶν ἰδιωμάτων*. Athens: Academy of Athens.
- LKN = Ίδρυμα Μανόλη Τριανταφυλλίδη, (1998). *Λεξικό της κοινῆς νεοελληνικῆς*, Thessaloniki: Ινστιτούτο Νεοελληνικῶν Σπουδῶν.
- LNE = Μπαμπινιώτης, Γ., (1998). *Λεξικό της νέας ελληνικῆς γλώσσας*. 2nd ed. 2002, Athens: Κέντρο Λεξικολογίας.
- LSJ = H. G. Liddell, R. Scott & H. S. Jones, *A Greek- English Lexicon, with a revised supplement*. Oxford: Clarendon Press, 1996.
- Vayakakos, D. (1977): *Γεώργιος Ν. Χατζιδάκις (1848-1941). Βίος και Έργον*. [Λεξικογραφικόν Δελτίον 2α, Παράρτημα]. Athens: Academy of Athens.

Modern Greek Dialectology and the Grammar of Medieval Greek

IO MANOLESSOU
Academy of Athens
manolessou@academyofathens.gr

0. Introduction

The Grammar of Medieval Greek is a five-year research project (2004-2009) whose aim is to describe in detail, for the first time, the Greek language between the years 1100-1700. As such, it is not a dialectological project *per se*, nor was its infrastructure initially geared towards the coverage of dialectal data; however, given its spatiotemporal coverage and the lack of previous systematic research in these domains, it has in fact become the largest project in Greek historical dialectology, as well as the largest digital repository of historical dialectal material. It is for this reason that this short presentation, which has no claims on originality, is included in the present volume (constant reference is made to various publications where the points touched upon here are discussed in detail).

In what follows, a short overview of the project (participants, aims, scope) will be given, followed by a description of its infrastructure (corpus of data, databases) and its contribution to Modern Greek dialectological research, on the basis of concrete examples.

1. The Grammar of Medieval Greek project

The Grammar of Medieval Greek project is located at the University of Cambridge, Faculty of Modern and Medieval Languages¹¹, and is funded by the Arts and Humanities Research Council. Its duration, as already mentioned, is 5 years, 2004-2009- this means that the basic research has already been finished and the results are in the process of being written up. By 2011 it is projected that these results will be published, by Cambridge University Press, in a volume entitled *A Reference Grammar of Medieval and Early Modern Greek*, with 6 authors, which are also the 6 members of the project: Professors David Holton and Geoffrey Horrocks (principal investigators) and Tina Lendari, Notis Toufexis, Marjolijne Janssen and Io Manolessou. The work of the project is assisted by an Advisory Board of international scholars, experts in the field of Medieval Greek texts, which convenes once a year.

The scope of the project is the description of the Greek language between 1100 and 1700, which covers roughly the time-period covered by the Dictionary of Medieval Greek Vernacular Literature of Emmanuel Kriaras (Kriaras 1967-). For this 6 century span, the aim is to describe the language of all Greek-speaking areas, i.e. from Italy to the West to Eastern Asia Minor in the East, Albania and Bulgaria in the North to Cyprus in the South. The project attempts to document this chronological and geographical variation, and track the spread and distribution of specific phenomena across these two parameters, time and space. This is the applicability of the project for the research on Modern Greek dialects: since, for the period under investigation, there is no standardized vernacular language, inevitably description is given by area. The end result is not "the Medieval Greek language" with local variants here and there, but a large diachronic and diatopic map of phonological, morphological and syntactic isoglosses.

2. Temporal coverage

The periodisation of the phase(s) of the Greek language between the end of the Koine and the Modern era is a notoriously difficult issue, and at times various scholars have proposed different dates both for its beginning (some placing in the 3rd and some in

¹¹ See www.mml.cam.ac.uk/Greek/grammarofmedievalgreek and Holton (forthcoming).

the 6th c.) and for its ending (some placing it in the 15th and some in the 17th c)¹². For the purposes of the project, the time-frame of Medieval Greek coincides with the limits defined by the lexicon of Kriaras (1967-) and therefore encompasses those texts written between the 11th and the 18th c., although notice is taken of developments dated close before and after those limits. The final product is therefore more accurately termed "Grammar of Medieval and Early Modern Greek", treating as it does a large corpus of texts from the 16th and 17th c. The reasons for the adoption of this time-frame are the following:

a) The aim to describe, as far as possible, the evolution of the Greek language during the medieval period. Therefore, primary weight is accorded to the type of texts which is usually termed "vernacular" i.e. close to spoken language and not imitating learned, archaising models. The abundant sources written in high registers during this period are only examined for comparative/corroborating purposes. Crucially, vernacular, "low-register" Greek texts become available only after the 12th c., with the rise of vernacular literature¹³.

b) The period prior to the 10th c. is, if not well, at least tolerably covered by the Grammar of non-literary papyri of the Roman and Byzantine periods (Gignac 1976-1981)- it is the language of the subsequent centuries for which no modern grammatical descriptions are available. In fact this was the main reason the project was proposed: the strange fact that although Greek is one of the most well-studied languages, both in its ancient form and in its modern one, for this specific period there simply does not exist any comprehensive grammatical/ linguistic description (Manolessou 2008, Holton & Manolessou 2010 and references therein).

c) the chosen time-period is a time of great linguistic variation, both chronological and geographical, which needs careful documentation and constitutes a necessary step for the investigation of the history of Modern Greek and the Modern Greek dialects. So in fact the project hopes to contribute not only to the research on Medieval Greek but also to the genesis of Modern Greek, constituting a necessary platform for the a future Historical Grammar of Modern Greek (this is another serious lack in the research on the Greek language).

3. Textual coverage

In order to ensure the reliability and authenticity of the data under examination, and to document, as fully as possible, the processes of spatiotemporal spread of the various phenomena, great importance has been accorded to the investigation of non-literary texts (notarial documents, wills, deeds of sale, marriage contracts, private letters etc.). The Grammar of Medieval Greek has prioritized research on such texts, for a variety of reasons: First of all, they are usually dated, named and of known provenance, so the linguistic information they provide is more precise than that of literary texts, which are usually of unknown author, and unspecifiable date and geographical origin. Secondly, they are usually transmitted in a single witness, never copied or copied only once, whereas literary texts in their present form are usually the result of several layers of consecutive copying which has distorted the linguistic picture, introducing features from different times and areas or deleting "outdated" features. Furthermore, non-literary texts are most frequently published as diplomatic editions, which provide an exact picture of the manuscript, spelling conventions, abbreviations etc. without any editorial interventions, while literary texts come to us through the intermediary of the editor, with all its

¹² For more details on the alternative periodisations of Medieval Greek, see Babiniotis (2002: 80-83), Holton & Manolessou (2010: 540-541).

¹³ For the notions "high" vs. "low" register in Medieval Greek see Toufexis (2008), and for the necessity relying on low-register texts for research on language change in Medieval Greek see Manolessou (2008).

advantages and shortcomings (corrections, normalization of orthography etc.)¹⁴. Thus the research conducted by this project is quite different from what has been done in the past concerning the investigation of medieval Greek, which was to a large extent based on the "classic" vernacular literary texts, like the Romances of Chivalry, *Digenis Akritis*, *Ptochoprodromos* etc..

However, the evidence of non-literary texts does not suffice to give a full picture of Medieval Greek, for three main reasons:

a) geographical distribution: Non-literary texts are not available from all Greek-speaking areas. Some areas are over-represented, with hundreds of available primary sources, and some are really under-represented, with almost no available texts at all. Here is a chart showing the geographical distribution of the collected texts so far. The areas under Frankish and Venetian occupation provide very rich legal archives, whereas the areas under Turkish occupation have very little to give. Crete is by far the best documented area, closely followed by the Heptanese and the Cyclades¹⁵.

c) chronological distribution: After the 15th c., non-literary evidence is abundant (mainly from the areas mentioned above). But from the 12th to the 15th non-literary documentation is very scarce. It comes mainly from two areas where a large number of monastery archives are preserved, S. Italy and Athos, which are of no great use concerning distribution, since it cannot be guaranteed that a text written or preserved there was in fact written by a native of the area.

d) genre limitations: some linguistic phenomena, especially in morphology, are very hard to come by in non-literary texts. For example, the 2nd person plural imperfect is extremely rare, especially in the passive- despite examining literally hundreds of texts, the collected attestations remain less than a dozen. First and second person verbal forms, unreal, counterfactual and future formations, genitive plural of feminine and neuter adjectives, are some of the forms that present the greatest difficulty during data collection, due to the nature of the available texts (mainly factual narratives and statements).

Nevertheless, the special attention accorded to texts of ascertainable local provenance has ensured that the Grammar of Medieval Greek is the basic source for anyone interested in the historical dialectology of later Greek: the first attestations of dialectal phenomena, the geographical distribution of linguistic features not common to all forms of Greek, as well as the spread of features which will ultimately form part of Standard Modern Greek can be investigated by using the corpus and the tools developed for the project. It is important to note that systematic research on the history of the Modern Greek dialects has never been undertaken before, and that, in comparison to most Modern languages, Greek is lagging seriously behind¹⁶.

4. The corpus

The textual corpus of the Grammar of Medieval Greek project consists of the following types:

A) .xml, .html, or .doc editions of texts, some, but not all, including apparatus criticus - ca. 2.500.000 words. This is an average-sized historical corpus¹⁷, created from the following sources: i) the TLG (*Thesaurus Linguae Graecae* - www.tlg.uci.edu), which is expanding towards the Medieval period and now contains several literary works (such as the *Chronicle of the Morea*, works of Cretan Drama, etc.) and, more importantly, non-

¹⁴ On the primary importance of datable and geographically localizable texts for linguistic research, and on the value and problems of non-literary sources for the investigation of Medieval Greek in particular, see Manolessou (2001), Manolessou (2008) and Markopoulos (2009).

¹⁵ A rough idea of the statistical distribution of sources by area is provided in Manolessou (2008b).

¹⁶ On historical dialectological research in Greece as compared to other countries see Manolessou (2008b).

¹⁷ For the notion "historical corpus", and the various historical corpora available for Greek as compared to those for other major languages, see Manolessou & Toufexis (forthcoming).

literary texts, such as the ca. 20 volumes of the Archives of Athos ii) other electronic versions of texts available through the Internet iii) texts donated by the modern editors or publishing houses themselves, in a generous gesture towards the project and iv) direct typing/transcribing from the printed editions by the members of the project. Scanning of printed texts and transformation to machine-readable form has been attempted only in a very small scale, as the results of OCR (optical character recognition) for polytonic texts of mixed linguistic form or not following standard orthography has proved unsatisfactory.

This section of the corpus is searchable: it can thus be used to locate specific phenomena, endings, collocations etc. However, because it is neither parsed nor tagged (a huge task which would require a research project on its own) there are limits on the types of searches than can be performed. For example, if one is interested in the "old" 3rd declension genitive inflectional suffix /os/ (e.g. τῆς γυναικός, τοῦ ἀνδρός, but later also τῆς κοπελός, τῆς Πατρός) there is no way to distinguish it from the 2nd declension nominative inflectional suffix /os/ (ὁ ἄνθρωπος, ὁ λόγος etc). Furthermore, because, as already mentioned, the non-literary texts exist in diplomatic editions, and because literary texts exist in many types of accentuation conventions (monotonic, Modern Greek polytonic, Ancient Greek polytonic), there is no unified spelling that one can use for electronic searches, and all alternative variants must be thought of (in this case, -ος, -ως, -οσ, -ωσ, -ὸς, -ὸς, -ὸς, -ὸς, -ὸς, -ὸς, -ὸς, -ὸς and so on...). Despite its limitations, the electronic corpus has proved invaluable in giving a rough idea of statistical frequency of the various phenomena and features.

B) .pdf texts of mostly non-literary texts, created from the following sources: i) downloading of out-of copyright publications from Anemi- the Digital Library of Modern Greek provided by the University of Crete (www.anemi.lib.uoc.gr) and from the Internet Archive (www.archive.org) and ii) photocopy and scanning of in-copyright printed publications. Because these files are image-file .pdfs, they are not searchable. However, they are useful for quick and remote access to publications, and for storing permanently and together indispensable sources for the history of Greek. In fact the project has created the largest archive of historical dialectal publications in Greece, with hundreds of digitized publications of local documents, many of them published in rare out of print periodicals or books.

C) .tiff and .gif images of medieval literary manuscripts. Thanks to the generous collaboration of an older project undertaken by the University of Sydney (under the direction of M. and E. Jeffreys), one of whose aims was to create microfilms of most major medieval Greek manuscripts, the Grammar of Medieval Greek is in possession of digitized image versions of several important manuscripts, indispensable for checking the validity of editions.

D) The project is also in possession of a considerable body of texts NOT in electronic form, i.e. just printed books or photocopies- a small library dedicated to research on Medieval Greek.

5. The database

The database of the Project can be divided in two parts. The first part is bibliographical: an effort has been made to create exhaustive bibliographies (searchable through keywords) on the following topics:

A) publications of literary texts. The aim is to create a Register of Authors and Works of the period under investigation, complementing that of the Kriaras Medieval Dictionary and providing: Standardised English language abbreviations for all works, information on the various alternative, old and new editions of texts, and basic information on each text (verse, genre, dating etc.).

B) Publications of non-literary texts: considerable effort has been expended in locating all available publications from the various areas of the Greek speaking world.

Using the Project's bibliography, one can easily locate all historical documents from a specific geographical area one is interested in¹⁸.

C) Linguistic publications pertaining to specific phenomena or areas of the period under investigation. Several of these are also available in scanned .pdf format.

The second, and most important part of the database, is the tool for excerpting phonological, morphological and syntactic phenomena. The project researchers go through the texts which constitute the corpus and excerpt chunks of text which are representative of a list of pre-determined phenomena and features. The project Database contains ca. 32.000 textual excerpts, all providing detailed linguistic and source information, and allowing comparisons as to date, place, lemma or linguistic category. It can be used to search for and group phenomena according to author, period, area, linguistic environment, grammatical category or lemma (lexical item). This electronic tool has facilitated and speeded up the project's research considerably. However, problems still remain:

a) **Skewed representation** (chronological and geographical). As discussed above, only areas under Frankish or Venetian control provide sufficient documentation- they heavily outweigh evidence from Turkish-occupied areas such as Macedonia, Thessaly etc. Similarly, the earliest centuries (11th, 12th) are almost exclusively represented by the two large archive depositories of Athos and S. Italy. Thus the absence of a phenomenon or feature from the corpus and the database does not by definition entail its absence from the area or period in question.

b) **obscurity of written sources**: Due to the very nature of historical linguistic investigation, which is based exclusively on written texts, linguistic information is frequently obscured by the nature of the written record. Thus, phonetic information is undetectable through the spelling (for example, it cannot be determined from the orthography whether the strong palatalisation, known as "tsitakismos" of velar consonants represented as <κ> or <τσ>, is in fact [c], [ts], [tʃ] or something else). Also, conservative spelling may conceal phonetic evolutions, such as the deletion of final /n/ or synizesis. Morphological and syntactic information is concealed through the conscious effort of authors/scribes to achieve a more archaic style or to avoid strongly characterized dialectal features.

c) **No statistical data available**. As already discussed, the corpus of the project does not lend itself to large-scale computerized searches (since the texts are not parsed or tagged, and several of them are in fact image files). And the texts constituting the corpus have been only partially excerpted, i.e. only representative samples have been recorded in the database- they are not entered word-by-word. Therefore, the information provided in the final grammar will be approximative only as far as frequency is concerned, with general descriptions such as "rare", "frequent", "absent so far from this type of text" etc.

6. A Sample of historical dialectological research

The following constitutes a sample of the work than can be done using the corpus, the tools and the methodology described above. The phenomenon in question, dental palatalisation, is dialectally restricted, and appeared at some point during the Medieval or Early Modern period.

[n] > [ɲ] and [l] > [ʎ] before [j, i]. The dental nasal /n/ and the lateral /l/ undergo palatalisation before front vowels and semivowels. One may distinguish palatalisation before [j], a semivowel resulting from synizesis of /i/, which occurs in all areas displaying the phenomenon of synizesis and is datable accordingly, and palatalisation before the front vowel /i/ which is dialectally restricted.

Evidence for the first type of palatalisation, before [j], is difficult to establish, since the Greek alphabet has no way of denoting it. However, because in Modern Greek it

¹⁸ The work done by the Research Centre for the History of Greek Law of the Academy of Athens has been an invaluable aid in the compilation of this bibliography.

appears in all dialects as well as in the Standard language (Newton 1972: 137), it must be comparatively early, perhaps simultaneous with the phenomenon of synizesis. Thus, spellings denoting synizesis through change of <ε> to <ι> or accent shift after /l/ and /n/ in all probability (but without certainty) already involve a palatalised [ʎ] or [ɲ]:

πλιόν BERG., *Apokopos* A 459 (Vejleskov)
τῆς ἐλιές / τῖς ἐλιές (1573, Ikaria, TSELIKAS 2000a: 2, 18.15)
λιόντα *Cypr. Canz.* 1.1 (Siapkaras-Pitsillidès)
νιούτσικε TRIV., *Ist. Re Skotsias* 270
τὴν νιότην [LIMEN.], *Than. Rod.* 466 trans. (Lendari)
βουνιά FALIER., *Thrin. Path. Stav.* 223
γονίους MACH., *Chron.* V 65.16 (Pieris/Nikolaou-Konnari)

Furthermore, in Crete and Cyprus the spelling <γν> and <γλ> is occasionally used in order to denote a palatalised sound, in imitation of Italian and French <gn> and Italian <gl> respectively, thus providing more direct evidence of the phenomenon:

ἀναμεγλιά TROILOS, *Rodol.* 2.465 (Aposkiti)
Ἐγλιάν / Ἐγλιάν (1679, Cyprus, PERDIKIS 1998: 16, 41.17)
περβογλοῦ (1699, Cyprus, PERDIKIS 1998: 39, 95.6)
καγλιοντουνα / καλλιόν νά *Cypr. Canz.* 53.8 app. cr. (Siapkaras-Pitsillidès)
ιστιν κεριγιαν / εἰς τὴν Κερύγνεια MACH., *Chron.* R 100.37
εγνια / ἐγνιά (1640, Cyprus, PERDIKIS 1998: 4, 11.9)
εἰς τα χρόγνια / εἰς τὰ χρόνια (1642, Crete, PAPADOPOULOS/FLORENDIS 1990: 21, 16.41)

Another graphematic indication of the existence of palatalised [ɲ] involves /m/ followed by the semivowel [j] in cases of synizesis: the combination [mj]+V results in [mɲ] + V, through consonantisation of the palatal semivowel to a palatal nasal. This outcome is frequently spelt <μνι> in Cretan literary and non-literary texts, and rarely in texts from the Cyclades and Cyprus:

ἐπεθυμνιάς FALIER., *Ist. On.* 75
ἐχλώμνιαεν *Thysia Avr.* 198 app. cr. (B)
μνιαν hora / μνιάν ὥρα CHORT., *Erof.* I. 408 trans. X (Legrand)
μνιά νύκτα / μιὰ νύκτα *Rim. kor.* (A) 2 app. cr. (Caracausi)
ἀσίμνια / ἀσήμια (1532, Crete, KAKLAMANIS/LAMBAKIS 2003: 152, 274.25)
τὰ κορμνιά μας (1549, Crete, MARMARELI/DRAKAKIS 2005: 9, 10.15)
νὰ ζημνιώση MACH., *Chron.* V 316.2 (Dawkins)
τοὺς Ρομνιοὺς / τοὺς Ρωμιοὺς (1614, Tinos, HOFMANN 1936: 1, 59.41)

Evidence for the second type of palatalisation, before [i], comes from the testimony of grammars of the period. Thus Girolamo Germano (PERNOT 1907: 51) and Simon Portius (MEYER 1889: 9-10, 88) state that in certain Greek-speaking areas the sounds [l] and [n] are pronounced like Italian <gli>, <gni>, giving the examples σώνει <σόgni>, and παρακαλεῖ <paracagli>. Unfortunately it is not specified which areas present the phenomenon (apart from Chios) but it is emphasised that it is a dialectal phenomenon best avoided.

Direct evidence of the phenomenon is provided by the spellings <gn>, <gn> in the Latin alphabet and <γλ>, <γν> in the Greek, which are quite frequent in Crete:

thegli / θέλει CHORT., *Erof.* I.379 trans. X (Legrand)
i angegli / οἱ ἀγγέλοι *Thysia Avr.* 6 trans. M
ossa bugli / ὠσάν πουλι *Pal. N. Diath.* 3421 app. cr.

na to pglithigni / νὰ τὸ πληθύνει FOSK., *Fort.* I.9 app. cr.

ftignia / φτήνεια (1653, Crete, PANOPOULOU 1991: 2, 429.21)

mu fagnisticchie / μοῦ 'φανίστηκε *Thysia Avr.* 593 app. cr. (M)

tu eogniu / τοῦ αἰωνίου (1661, Chandax / Crete, MAVROMATIS 1986: 10, 102.1)

ci gnictes / τσι νύκτες CHORT., *Erof.* I.403 (Legrand)

narghignisso / ν' ἀρχινήσω FOSK., *Fort.* I.97 trans.

For other, published examples of historical dialectological work conducted with the means provided by the Grammar of Medieval Greek project, see Manolessou & Toufexis (2009), which discusses the phenomenon of the change of /l/ to /r/ and vice versa in Medieval and Modern Greek and Manolessou (2010), which provides a detailed account of the Medieval form of the Cypriot verb system.

7. Conclusion

The Grammar of Medieval Greek project has been the means of creating several primary tools for the investigation of the history of later Greek: bibliographies, textual corpora, and databases containing annotated examples. This material is currently being used in order to compile the *Reference Grammar of Medieval and Early Modern Greek*, but it is hoped that it will also be used in the future for research on the history and analysis of the Modern Greek dialects.

References

- Babiniotis, G. (2002): *Συνοπτική ιστορία της ελληνικής γλώσσας*. 5th edition, Athens.
- Gignac, F. T. (1976-1981). *A Grammar of the Papyri of the Roman and Byzantine Periods*. Milano: Cisalpino-Goliardica.
- Holton, D. (forthcoming): "The Cambridge Grammar of Medieval Greek Project: Aims, Scope, Research Questions". In G. Mavromatis, ed., *Neograeca Medii Aevi VI*, Ioannina, October 2005.
- Holton, D. & I. Manolessou (2010): "Medieval and Modern Greek". In E. Bakker (ed.), *A Companion to the Ancient Greek Language*, Oxford: Wiley/Blackwell, 539-563.
- Kriaras, J. (ed.) (1967-). *Λεξικό της Μεσαιωνικής δημόδους γραμματείας*, 1100-1669, vol. 1-15. Thessaloniki.
- Manolessou, I. (2003): "Οι μη λογοτεχνικές πηγές ως μαρτυρίες για τη γλώσσα της μεσαιωνικής περιόδου". *Λεξικογραφικόν Δελτίον* 24: 61-88.
- Manolessou, I. (2008): "On Historical Linguistics, Linguistic Variation and Medieval Greek". *BMGS* 32: 63-79.
- Manolessou, I. (2008b) «Μεσαιωνική γραμματική και ιστορική διαλεκτολογία: παρατηρήσεις με αφετηρία την κυπριακή διάλεκτο». Στο *Πρακτικά του Ε' Διεθνούς Συνεδρίου Νεοελληνικής Διαλεκτολογίας*, Αθήνα: Ακαδημία Αθηνών, 425-448.
- Manolessou, I. (2010) "Κυπριακή γραμματική Γ': Ρήματα". Στο Α.Βοσκόσ, Δ. Γούτσος & Α. Μόζερ (επιμ.), *Η ελληνική γλώσσα στην Κύπρο από την αρχαιότητα ως σήμερα*, Αθήνα: Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, 68-87.
- Manolessou, I. & N. Toufexis (2009): "Phonetic Change in Medieval Greek: focus on liquid interchange". In *Proceedings of the 8th International Conference on Greek Linguistics, Ioannina, August 30th- September 2nd 2007* [e-book], 291-306.
- Manolessou, I. & N. Toufexis (forthcoming): "Corpus Linguistics in historical dialectology: a case-study of Cypriot". In *Proceedings of the 3rd International Conference on Modern Greek Linguistics and Linguistic Theory*, Nicosia.
- Markopoulos, Th. (2009): "Η έρευνα της γλωσσικής αλλαγής στα κείμενα της Μεσαιωνικής Ελληνικής: μεθοδολογική διερεύνηση των πηγών". In *Proceedings of the 8th International Conference on Greek Linguistics, Ioannina, August 30th- September 2nd 2007* [e-book], 987-998.
- Meyer, W. (ed.) (1889): *Simon, Portius, Grammatica Linguae Graecae Vulgaris*, 1638. Paris: Vieweg.
- Newton, B. D. (1972): *The Generative Interpretation of Dialect*. Cambridge: CUP.
- Toufexis, N. (2008): "Diglossia and Register Variation in Medieval Greek." *BMGS* 32: 203-19.
- Pernot, H. (ed.) (1907): *Girolamo Germano, Grammaire et vocabulaire du grec vulgaire, publiés d'après l'édition de 1622*. Fontenay-sous-Bois (Seine).

Performing dialectal talk: Differentiating gender roles via direct speech representation*

ARGIRIS ARCHAKIS,
University of Patras
archakis@upatras.gr

SOFIA LAMPROPOULOU
Lancaster University

DIMITRIS PAPAACHARIOU
University of Patras
papaz@upatras.gr

Abstract

In this paper we argue that dialectal variation should not only be approached as a geographical reflex, but also as an interactional resource for various communicative objectives, in line with recent discourse analytic approaches. Our analysis concentrates on the performative strategies via which an old dialectophone stylizes her father and her mother in a constructed story. We argue that our informant, by manipulating the narratives she produces in terms of both dialectal and discursal features, she adopts the discourse identity of performer.

1. Introduction

Research within the framework of dialect geography and traditional dialectology was mainly based on the assumption that region acts as the cause for a particular kind of linguistic variation called dialectal variation. This assumption includes the beliefs that dialects are spoken by homogeneous, non-mobile and often rural social groups living in a situation of communicative isolation within a particular region. However, in our modern, or post-modern, world, we scarcely meet this sort of homogeneous and stable local groupings. Rather, the contemporary world is characterized by heterogeneous communities consisting of mobile people who spend periods of their life in different places and who quite often change occupations and life styles (see Johnstone 1999: 506-507, 515).

Taking into account this new diverse population composition in contemporary country-sides, modern approaches to dialect analysis are not constrained to pose research questions of the type what a dialect is, i.e. what are its defining and differentiating features in all or some levels of linguistic analysis, but, from a discourse analytic perspective, they are also interested in how dialectal features can be used so that a bidialectal speaker can attain various communicative goals in various contexts of communication. This means that regional dialectal differences are not so much approached as situational reflexes, but also as indices of symbolic values, being one of the speakers' strategic means for activating meaning potential relevant at different points of their interactions (Rickford & Eckert 2001: 4-6, Coupland 2001: 209).

In this paper our aim is to discuss certain performative functions of a northern Greek dialect, namely the Lesbian Dialect. For this purpose we have chosen to analyse one conversational narrative produced by an old Greek woman throughout her conversation with a researcher (see also Archakis et al 2009). The Greek woman, whose name from now on will be Matoula, was an immigrant for more than twenty years in Athens and has returned back home at the village, Afalonas, in the island of Lesbos. She had accepted to talk her dialect and about her dialect with the researcher. We will analyse the instances of dialectal features she produced in the selected story, mainly identified within narrative direct speech. We will show that her switching from the standard Modern Greek to the production of dialectal features is closely related to the discourse identity she adopts.

* We would like to thank Prof. Ralli for her support and for the access she gave us to the recorded data, which are product of the research project *Documentation and Description of the Dialect of Eastern Lesbos. Comparison with the Asia Minor Dialect of Aivali and Moschonisia* and part of the *Greek Dialects' Corpus*.

More particularly, we will show that when she is asked to use the dialect, she does not code switch and speak it, but acting as a performer she presents it under specific conditions. Special attention will be given to the way she stylizes two main characters in the story under investigation, namely her father and her mother.

2. Key – concepts

2.1. Time, place and direct speech in conversational narratives

As we have already pointed out, in the narrative under consideration the dialectal features are identified in direct speech instances. However, it is not in every narrative with direct speech produced by Matoula, that we find dialectal features. A crucial presupposition is that the time and place of the represented sequence of events should belong to the remote past, when her parents and their relatives and neighbors were alive. After all, it is the voice of these people that Matoula animates in order to present the dialect. Thus, in what follows we will elaborate on these basic concepts of narrative analysis, i.e. *time, place and direct speech*.

Narrowing down our focus on physical setting we introduce a distinction, proposed by cultural geographers, between *physical spaces* and *cultural places*. Johnstone (1999: 516) points out that “[p]hysically delimited areas -spaces- are not places unless they have meaning for people as distinct from other places”. As we shall see in the analysis of the selected narrative that follows, it seems that Matoula perceives the dialect as part and parcel of a foretime cultural place when it was inhabited by people with a different value system than the current one. Thus, the dialect can only be spoken by- and through-them and not directly by her.

In connection with the roles of *narrative time* and *place* in the development of a story, Georgakopoulou (2003: 415) stresses their exploitation as interactive resources in the here and now of a storytelling situation. She explains that different places in their interaction with time “create affordances” for “different sets of expectations about what sorts of action and interaction with what sorts of participants can take place where and when” (ibid: 424). Based on this observation, we shall show that Matoula discursively constructs and locates the narrative time and place in the remote past so that her story participants can afford speaking the dialect.

If, as we maintained, the dialect, according to Matoula’s practice, can mainly be spoken by people of a cultural past, then a possible way for Matoula to comply with the request of the researcher and speak the dialect is to animate the voices of these people via *direct speech*. Thus, our focus is placed on direct speech, namely the report of voices that were uttered in anterior context from the current one. We could point out that direct speech, due to its grammatical characteristics, gives the impression of a verbatim reenactment of the original utterance (Holt 2000). In view of this property, direct speech, appearing mostly at the climax points of narratives, can be seen as an internal evaluative device (Labov 1972) contributing to the vividness and dramatization of the reporting utterances and to interpersonal involvement (see among others Tannen 1989).

2.2. Discourse identity

Our analysis draws upon a dynamic approach to identity construction. According to this approach, identities are not static and stable properties that reside in peoples’ minds but emerge through discourse, where they are dynamically recreated. To this end, people project different aspects of their identities, depending on different contexts on the basis of various and different forms of verbal behaviour (Antaki and Widdicombe 1998).

In our attempt to trace Matoula’s shifts from the Modern Greek Koine to the dialect and vice-versa, we will apply the concept of *discourse identity* proposed by Zimmerman (1998), who treats identity as “an element of context for the talk-in-interaction” (ibid: 87). According to Zimmerman, discourse identities emerge from the sequential organization of talk and “are integral to the moment-by-moment organisation of the interaction” (ibid:

90). Thus, participants assume discourse identities which orient them to certain activity types and their respective interactional roles within them, such as the roles of current speaker, listener, story teller, story recipient, questioner, and answerer.

In the analysis of the selected narrative that follows, we will claim that Matoula adopts a specific discourse identity, namely that of the performer, in order to assign different voices to the represented characters of her stories, i.e. her father and mother, who are the foretime inhabitants of her village and presumed authentic speakers of the dialect. According to Bauman (1986: 3), *performance* is a mode of communication that highlights “the way in which communication is carried out, above and beyond its referential content”. Georgakopoulou (1997: 144), concentrating in particular on the main purpose of Greek performances, maintains that it purports to “create an immediate, empathetic narration. It is by staging a multi-media show (auditory and visual) that storytellers aim at creating an internal emotional connection with the narration and the audience”.

What is important for our study is to elaborate on the “auditory element” that accompanies the direct speech sequences, i.e. the main loci of dialectal talk production in our data. More particularly, we will demonstrate how Matoula performs, rather than speaks, the dialect. Based on Rampton’s (1995) notion of *crossing*, special attention will be given to Matoula’s *crossing practices*, that is to the fact that she selects the appropriate dialectal features in order, not only to construct herself as capable of switching from Koine to the dialect and vice versa, but also in order to assign different identities to the represented voices. In other words, we will show how she performs different dialectal and prosodic features in order to represent the voices of her father and her mother, *styling* them in different ways (Rampton 1999).

3. The data of the study

The recording of Matoula’s conversation was part of a bigger project, under the supervision of Prof. Ralli, aiming to record and analyse the dialect of Eastern Lesbos.¹⁹ This particular informant, having lived for more than twenty years at the capital of Greece and being capable of using the Standard Modern Greek, presented a very interesting behavioural and speech pattern. In particular, she communicated with the field-worker mainly in Standard Modern Greek Koine, although she knew beforehand that the researcher was interested in recording the dialect. To a straightforward request from the field-worker to speak the dialect, she replied that she would do so only in particular contexts.

Careful study of her recorded dialogues with the field-worker reveals that Matoula’s dialectal talk mainly lies in 11 occurring conversational narratives that were inspired by topics referring to the cultural past of her village. More specifically, dialectal features appear in the 80 direct speech instances that are identified within these narratives. In this paper, we will particularly concentrate on the analysis of the performative strategies through which Matoula stylizes her father and her mother in a selected story.

4. Analysis

In the narrative episode under examination we will demonstrate a recurrent shift in the presentational mode of Matoula’s stories. According to Bauman (1986: 66), “there is a need for ways of marking the difference between the voice of the narrator in the present storytelling context and the reported speech of the actors in the original event being reported”. Matoula systematically distinguishes the way she recounts circumstances and actions from the way she replays interactions. In particular, the diegesis mode is carried

¹⁹The name of the research project is *Documentation and Description of the Dialect of Easter Lesbos. Comparison with the Asia Minor Dialect of Aivali and Moschonisia* which is funded by the EU and the Greek Ministry of Education (Program EPEAEK-PYTHAGORAS), under the supervision of Prof. Ralli. The recorded material became part of the *Greek Dialects’ Corpus*, which is hosted at the Linguistics Lab of Modern Greek dialects at the University of Patras, Greece.

out mainly in the standard Modern Greek Koine, whereas the mimesis mode draws upon dialectal recourses. It is through this shift from telling to showing and reenacting that Matoula acquires the discourse identity of the performer.

In order to underpin this claim, we will provide one narrative extract where Matoula represents events of a past spatiotemporal context, including the representation of utterances that were produced in anterior context. These are events that deal with the everyday life of her family in the village when Matoula was very young, i.e. approximately sixty years ago (in relation to the time of the recording).

The following episode describes a habitual event that occurred when Matoula's father came back home after work (see also Archakis *et al.* 2009). In particular, it includes a small quarrel he occasionally had with her mother concerning lunch, as part of the father's bad mood due to hard work.

1. MT: *Τη διάλεκτο να σου πω τώρα πώς μιλάγαμε στο σπίτι τον καιρό που ζούσε η μάνα μ η γιαγιά μου πριν και πριν μάθουμε τα γράμματα τελοσπάντων και αυτά ε;*²⁰
As for the dialect, I will now tell you how we were speaking at home when my mom and my grandmom were alive and before, before we learnt how to read and write and so on

2. Fw: Αυτό ακριβώς
 Exactly

3. MT: *Ναι. Ε να σας πω μόλις ερχόταν ο πατέρας μ και ήταν λίγο θυμωμένος άρχιζε*
 1a.²¹ **Μωρή Φθυμίγια** ((γέλιο)) **που είσαι μωρή τσι σ' έχασα;** Η μάνα μ ήταν η Φθυμίγια
 [mo'ri fθi'mi:ja:: ... ((laugh)) 'pu 'ise mo'ri tsi 'sexasa?]
 Yes, let me tell you, when my dad was coming home and he was a bit angry he was starting
 1a. **mori**²² **Fthimigia** ((laughing)) **where have you been and I've lost you?**
 Fthimigia was my mom

4. Fw: Μμ
 Hmm

5. MT: 2b. **Ναι Γιάννη, έδγιω είμαι, ούι, ούι, τι κάνς;**
 ['ne jani'eðjo ime 'ui 'ui 'ti 'ka:ns?]
Yianni I'm here no no how are you?

3a. **καλά, εσύ τι γίνεσαι. Ε τι φαγί έκανες σήμερα;** Λέει
 [ka'la: e'si ti 'jinese.. E: 'ti fa'ji 'ekanes 'simeras?]
Fine, and you? Um what kind of food have you made for today she says

4b. **φασούλες.**
 [fa'sules]
Beans

5a. **Πάλι φασούλες λεγ' θα φάμε; Άντε μωρ' τσε δε μπορώ να τρώγ' όλ μέρα**

²⁰ Words in italics reveal the setting of the story. The location is the Afalonas village and the time is approximately fifty years ago.

²¹ Direct speech instances appear in bold and are numbered. In order to facilitate the tagging of the turn-taking instances, we include a letter which stands for a different represented voice, just after the serial numbers: *a* stands for the father's voice and *b* for the mother.

According to Prof. Ralli, who is a native speaker of the Lesbian dialect, Matoula's direct speech instances are not representative of the system of the Lesbian Dialect. There are cases where dialectal features appear even in environments where they shouldn't appear and cases which are neither dialectal nor of Standard Modern Greek. Due to these inconsistencies, Matoula's talk, although including various and different dialectal features, cannot be considered as representative of a Northern Greek dialect. This observation corroborates our approach relating to the performative aspect of her speech.

²² *Mori* is an untranslatable Greek discourse marker that signals intimacy.

*Performing dialectal talk:
Differentiating gender roles via direct speech representation*

φασούλς.

[ˈpali faˈsules θa ˈfame?] [ˈade ˈmor tse ðe boˈro na ˈtroj ˈol
ˈmera faˈsuːls]

will we have beans again? I can't be eating beans all days.

Τον πιάναν και τα νεύρα. **Να πάρ η διάλογος φασούλς τσι σένα κι το κεφάλ σ.**
[na ˈpar i ˈðaolos s faˈsuls tsi ˈsena ki to ceˈfal s]

He was starting pissing off. **Damn the beans and you and your head.**

6b. **Αχ τι πάθαμ.**

[ˈax ti ˈpaθam]

Oh my goodness.

7a. **Θα πά να φάου έξω ((γέλιο))**

[θa ˈpa na ˈfau ˈeksɔ]

I'll eat out ((laughing))

8b. **ε που θα πας βρε άθρουπε, έχεις φαγί, μα έχς παράδες για θα πας να φας έξ**

((γέλιο))

[e ˈpu θa ˈpas vre ˈaθrupe .. ˈeçis faˈji ma . ˈexs parˈades ˈja θa ˈpas na
ˈfas ˈeks]

but where will you go man, there is food, and do you have money in order to go out? ((laughing))

λοιπόν γινόταν ένα καβγαδάκι τελικά

Well there was a small quarrel and finally

9a. **άντε φασουλάδα μι ελιές είναι μια χαρά. Φάγαν τα μωρά;**

[ˈade fasuˈlaða mi eˈles ˈine ˈmja xaˈra .. ˈfayan ta moˈra?]

ok beans with olives are fine did the kids eat?

10b. **Φάγαν.** τα παιρνε.

[ˈfayan]

They did, he took them

11a. **Για έλα βρε μορέλι, έφερα τσι μιντούδς, τσι καραμιλούδς, τσιπραματέλια να φάτε.**

[ˈja ˈela vre moˈreli ˈefera tsi miˈnduðis tsi karamiˈluðis, tsi
praɣmaˈteɫa na ˈfate]

come here my baby I've brought candies and stuff to eat.

Τρώγαμε τον φλούσαμε τον πατέρα, τον αγκαλιάζαμε, τα κάναμε όλα αυτά.

We were eating we were kissing dad we were hugging we did all these.

What is particularly interesting in the above story is the fact that the information on orientation, i.e. the information related to the physical and temporal setting of the story is mainly produced in Koine (turn 1: *Τη διάλεκτο να σου πω τώρα πώς μιλάγαμε στο σπίτι (...)* ["as for the dialect, I will now tell you how we were speaking at home"]), whereas the direct speech quotations are produced by Matoula using a lot of dialectal features (lines 1a-11a).

In what follows, we will show that she employs different dialectal features in the direct speech instances of the narrative episode under examination in order to represent the voices of her father and her mother.

In terms of dialect use, we observe that both genders employ a variety of dialectal features throughout all the narratives. Nevertheless, Matoula seems to allocate different dialectal features to her father's voice and different to her mother's mimicry. In particular, we focus on the dialectal features of a) raising of the unstressed middle vowels, b) the

deletion of the unstressed high vowels, c) tsitacism, d) local diminutive suffixes, as well as e) local masculine article.

As to the raising of the unstressed middle vowels /e/ and /o/ to [i] and [u] in father's voice, according to table 1, in thirty eight possible locations of raising –marked in italics-, we observe fourteen raised realizations –marked in bold letters:

Table 1: *Raising of unstressed middle vowels in father's voice*

Possible locations of raising	Actual realizations of raising
1a. [mo'ri, 'ise, mo'ri, tsi]	1a. [tsi]
3a. [e'si, 'jinese, ekanes, 'simera?]	5a. [tsi, tsi ci]
5a. [fa'sules, 'fame? 'ade, 'mor, tsi, bo'ro, 'ðaolos tsi ci to ce'fal]	7a. ['fau,]
7a. ['fau, 'ekso]	9a. [fasu'laða, mi]
9a. ['ade, fasu'laða, mi, e'les, 'ine, mo'ra?]	11a. [mu'reli, tsi, mi'nduðis, tsi, karami'luðis, tsi,]
11a. [mu'reli, 'efera, tsi, mi'nduðis, tsi, karami'luðis, tsi, 'fate]	

On the other hand, in mother's voice, there is only one occasion of a raised middle vowel, out of seven possible ones, as can be observed in table 2.

Table 2: *Raising of unstressed middle vowels in mother's voice*

Possible locations of raising	Actual realizations of raising
2b. [eðjo ime]	8b. ['aθrupe]
4b. [fa'sules]	
8b. [vre, 'aθrupe, par'aðes]	

In relation to the deletion of the unstressed high vowels /i/ and /u/ in father's voice, in the line 5a we can see five possible locations for deletion (in table 3), where actually the deletion did occur. Interestingly, vowel deletion occurred in two more cases, where the vowel is not a high, but a middle one, indicated by bold empty brackets in table 3.

Table 3: *Deletion of unstressed high vowels in father's voice*

Deleted unstressed high vowels	Deleted unstressed middle vowels
5a. ['mor[], 'ol[], 'par[], s[], ce'fal[]s]	5a. [fa'su:l[]s, fa'sul[]s]

On the other hand, in mother's speech, high vowel deletion appears in two out of four possible locations (indicated by brackets in table 4)

Table 4: *Deletion of unstressed high vowels in mother's voice*

Deleted unstressed high vowels	Undeleted unstressed high vowels
2b. 'ka:n[]s	2b. ['jan[i]]
8b. 'ex[]	8b. ['eç[i]s]

Another phonological phenomenon of the dialect of Lesbos is the transformation of the palatal voiceless obstruent [c] to +delayed release [ts], before front vowel [i], known also as tsitakism. Interestingly, in father's voice, tsitakism appears almost in every possible case -that is in seven out of eight instances- (see table 5), but not even once in mother's voice, not only in this particular extract, but in the other narratives as well.

Table 5: *Tsitakism in father's voice*

Possible locations of tsitakism	Actual realizations of tsitakism
1a. [tsi]	1a. [tsi]
5a. [tsi, ts, tsi, ci]	5a. [tsi, ts, tsi]
11a. [tsi, tsi, tsi]	11a. [tsi, tsi, tsi]

The same pattern appears with the other two morphological dialectal features, that is a) the local diminutives suffixes /-eli/, /-elia/ and /-uði/, /uðes/, and b) the masculine article /i/. In particular, we can find the local variants in father's voice (table 6, in bold), but nowhere in mother's speech.

Table 6: *Morphological dialectal features in father's voice*

Local diminutives suffixes	The masculine article as [i]
11a. [ja 'ela vre mo'reʎi 'efera tsi mi'nduðes tsi karami'luðes, tsi prayma'teʎa na 'fate]	5a. [na 'par i 'ðaolos s fa'suls tsi 'sena ci to ce'fal s]

Based on the above observations, we could argue that Matoula seems to assign different identities to the different voices she presents as speaking. In particular, the father is presented as employing more frequently the features that have been stereotypically associated with the local dialect. On the other hand, the mother, when she is represented to talk to her husband, does not seem to employ the stereotypical dialectal features that frequently; actually, in some cases she does not employ them at all. Taking into consideration the constructed nature of direct speech, namely that the narrator-animator eventually reserves for herself the authorial and principal rights (Goffman 1981), and, thus, direct speech representation is constructed on the basis of narrator's communicative goals, we could argue that Matoula constructs her father and her mother in different ways. Gender identity seems to play an important role in this construction.

In particular, it seems that identities like masculinity and femininity come into play and are related to stereotypical assumptions concerning gender and dialect. Drawing on Rampton's remarks (1999: 421), we could point out that Matoula uses dialectal features in the discursive practices of direct speech to appropriate and reproduce influential images and stereotypes of gender groups that she does not herself straightforwardly belong to. More specifically, the way Matoula stylizes her father with regionally marked variety deviates from the linguistic norm that, in this case, is Koine. On the other hand, the mother seems to be represented much more aligned with the norm, as the ellipsis of many dialectal features result in her speech approaching Koine and, in a way, distancing herself from the dialect (see Georgakopoulou 2005: 175). The association of femininity with normative linguistic behavior and of masculinity with more deviant, in a way, linguistic behavior is a sociolinguistic pattern that has been observed by many researchers (see Trudgill 1974, Labov 1990). On this basis, we assume that through her discursive constructions, Matoula seems to exploit the sociolinguistic stereotypes in order to construct contrastively gender identities.

5. Discussion and concluding remarks

In this paper, we have presented a narrative episode produced by a dialectophone who, throughout the conversation with the researcher, switched between the use of Koine and the production of dialectal features. We pointed out that Matoula considers the dialect as part and parcel of a distant cultural past. Thus, she speaks it out mainly through the performance of voices that belong to this past. The main vehicle for this performance is direct speech representation. We therefore argued that Matoula, when representing the voices of other people, extracts the dialect from a distant past, as if she performs a role in a play. To this end, Matoula adopts the discourse identity of the performer.

Particular emphasis was put to the fact that Matoula, by adopting the discourse identity of the performer, is able to assign different identities to the represented voices. In a selected story we observed that the identities attributed to her father and mother seem to correspond to gender stereotypes. We therefore argued that the (represented) dialect may consist of a vehicle that reflects and sustains socio-cultural values and (stereotyped) assumptions regarding gender. Following Rampton's claims (1999: 423), instead of simply concentrating on the functioning of the dialect of Afalonas within a context, we paid particular attention to "the complex (...) sociolinguistic processes involved in moving it **across** from one context to another".

Within this framework of analysis, our main finding lies in the fact that the old dialectophone uses the dialectal features on the precondition that she guides her audience

to her socio-cultural past. For this purpose, she discursively constructs this socio-cultural past and its foretime characters. In this way, she creates the affordances for the dialect to be spoken. It is the people who belong to this past that, according to her discursive practice, have the right to speak the dialect, at least in front of an out-group. Thus she is able to invoke and index their tradition and its symbolic value that is attempted to be preserved in the current geographical space of her village.

References

- Antaki C. & Widdicombe S. (1998). Identity as an achievement and as a tool. In Charles Antaki and Sue Widdicombe (eds.) *Identities in Talk*. London: Sage. 1-14.
- Archakis A., Lampropoulou S. & D. Papazachariou (2009). 'I Will Use the Dialect Only in Specific Contexts': Performing the Dialect through Direct Speech Representation. Evidence from a Northern Greek Dialectophone. *Journal of Greek Linguistics* 9: 5-33.
- Bauman R. (1986). *Story, Performance and Event. Contextual Studies of Oral Narrative*. Cambridge: Cambridge University Press.
- Coupland N. (2001). Language, situation, and the relational self: Theorizing dialect-style in sociolinguistics. In Penelope Eckert & John R. Rickford (eds.) *Style and Sociolinguistic Variation*. Cambridge, U.K.: Cambridge University Press, 185-210.
- Georgakopoulou A. (1997). *Narrative Performances: A Study of Modern Greek Storytelling*. Amsterdam, The Netherlands /Philadelphia, Pennsylvania: John Benjamins.
- Georgakopoulou A. (2003). Plotting the "right place" and the "right time": Place and time as interactional resources in narrative. *Narrative Inquiry* 13(2): 413-432.
- Georgakopoulou A. (2005). Styling men and masculinities: Interactional and identity aspects at work. *Language in Society* 34: 163-184.
- Goffman E. (1981). *Forms of Talk*. Oxford: Blackwell.
- Holt E. (2000). Reporting and reacting: Concurrent responses to reported speech. *Research on Language and Social Interaction* 33: 425-454.
- Johnstone B. (1999). Uses of southern-sounding speech by contemporary Texas women. *Journal of Sociolinguistics* 3(4): 505-522.
- Labov W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov W. (1990). The interaction of sex and social class in the course of linguistic change. In *Language Variation and Change*. 2: 205-254
- Rampton B. (1995). *Crossing: Language and Ethnicity among Adolescents*. London/New York: Longman.
- Rampton B. (1999). Styling the Other: Introduction. *Journal of Sociolinguistics* 3(4): 421-427.
- Rickford J. & Eckert P. (2001). Introduction. In Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation*. Cambridge, U.K.: Cambridge University Press, 1-18.
- Tannen D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. New York: Cambridge University Press.
- Trudgill P. (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Zimmerman D. (1998). Identity, context and interaction. In Charles Antaki and Sue Widdicombe (eds.) *Identities in Talk*. London: Sage. 87-106.

The perception of plosive gemination in Cypriot Greek

SPYROS ARMOSTI

Phonetics Laboratory, University of Cambridge

armostis@gmail.com

1. Purpose of the study

Cypriot Greek plosive and affricate geminates are rather unusual, as they are distinguished from singletons by both longer closure (i.e. the universal cue to gemination) and longer release. This situation caused a recent debate in the literature on which of the two is the main cue to gemination of Cypriot Greek plosives and affricates. Arvaniti and Tserdanelis (2000) argued that the duration of the closure (CD) serves as a salient cue to gemination (as is the case with other geminating languages, such as Italian and Turkish), while aspiration (Asp) was considered to be an enhancing cue (p. 562). Similarly, Muller (2001) regarded CD as the primary correlate to gemination in Cypriot Greek, with aspiration as a secondary cue, albeit a very important one. Botinis et al. (2004) also concluded that gemination is mainly achieved by an increase in CD, combined with an increase in aspiration. Christodoulou (2007), on the other hand, suggested that aspiration is the main cue to gemination of stops in Cypriot Greek, while closure duration is a secondary cue; this claim is based on the fact that the closure is inaudible phrase-initially, yet minimal pairs of word-initial geminate stops are robustly distinguished in phrase-initial position.

Most of the aforementioned studies were acoustic studies, therefore the cues to which they referred were *acoustic* cues; if one is to make any claims about the *perception* of those cues, this should only be done based on a perceptual study. Botinis et al. did conduct a perceptual study (along with their main acoustic one), which showed that aspiration was a robust perceptual cue to gemination in utterance-initial position. In this specific case, the absence of the CD cue (since it is unperceivable utterance-initially) makes aspiration the only audible cue to gemination; but this is not a definite proof that, in cases where the duration of the closure is in fact perceivable (like word-internally), aspiration is still the *main* perceptual cue to gemination. Actually, words with initial geminates are far fewer than words with internal geminates, a fact that makes the scenario of word-initial geminates appearing in phrase-initial positions unusual. Thus, in the vast majority of the occurrences of geminate stops, the closure is actually perceivable.

Muller did test the perception of word-initial gemination both in utterance-initial position and intervocalically, but her results should be treated with caution due to a fatal flaw in the design of the experiment regarding the stimuli selected. In her experiment the subjects heard the first syllable plus the lateral of the words /te'llerɔn/ 'frame' and /tɛ'lljɛzzɔ/ 'I wire' (among others) and had to discern whether the syllable came from the one or the other word. Those test syllables were heard in utterance-initial position and in utterance-medial position, and it was hypothesised that in utterance-medial position (where both CD and Asp are present) the gemination contrast would be better perceived than in utterance-initial position (where the CD cue is perceptually absent). The results showed that there was no difference between the two positions, as in both cases the discrimination between singletons and geminates was very clear (something that implies that the presence or absence of the CD cue does not make a difference). However, the lateral in the second word undergoes palatalisation, hence the two test syllables [tɛl:] and [tʰ:ɛʎ:] did not differ only with regard to their stop, but also with regard to the lateral. Thus, the listeners could distinguish the two syllables aided by an additional cue, that of the palatal lateral. This shortcoming means that the results regarding the perception of coronal gemination (and the perceptual primacy of Asp over CD) are not reliable.

A better way for a study to compare the perceptual weight of the closure vs. aspiration would be to test the target stops in word-internal position (so as to ensure that both cues are present) and to manipulate the duration of their closure and aspiration in order to explore the difference in perception that this durational alteration would cause (something that was not done in the previous studies). Such a perceptual study is the subject of this paper. In the sections that will follow, the design and conduct of the perceptual experiment will be presented, followed by the results and their analysis, which will lead to the discussion of the main question of this paper, namely what the relative perceptual importance of the various cues to gemination of Cypriot Greek stops is.

2. Method

In order to test the perceptual correlates of the gemination of word-internal stops, tokens of such stops were manipulated to produce stimuli with various durations of their closure and aspiration; these stimuli were then listened to by native speakers of Cypriot Greek, who were asked to respond whether they heard a singleton or a geminate stop.

The creation of the stimuli, the procedure of the test, and the analysis to which the results of the study were subjected, are presented in this methodological section.

2.1. Material

For the creation of the stimuli of the perceptual study, tokens of minimal pairs which differed in containing word-internal singleton or geminate stops were recorded. The method of recording these test tokens will be presented first.

2.1.1. Recording the test tokens

2.1.1.1. Test sentences

Two words forming a minimal pair based on the quantity of the alveolar stop they contained served as tokens for the recordings. The reason for the selection of only one place of articulation was to confine the resulting stimuli to a reasonable number, in order for the perceptual study not to become particularly long in duration, and hence tiring for listeners. The alveolar place of articulation was chosen as an intermediate point between the extremes of the labial and velar place. The tokens for the recording were the words [ˈpitɛ] (i.e. ‘hose’ imperative, singular) and [ˈpitʰːɛ] (i.e. ‘pie’). The two tokens were embedded in a carrier phase as shown in Table 1.

Table 1: *The two test sentences.*

Test sentence 1	Εν είπα «πίττα», είπα «ράντισε».				
	en	ipe	pite	ipe	ndise
		l	s	h	s
	EG	aid-I	ose	aid-I	ray
I didn't say 'hose', I said 'spray'.					
Test sentence 2	Εν είπα «πίττα», είπα «κκέικ».				
	en	ipe	pitʰːɛ	ipe	eik
		l	s	p	s
	EG	aid-I	ie	aid-I	ke
I didn't say 'pie', I said 'cake'.					

2.1.1.2. Speaker

The speaker recorded was EE, a female speaker of ‘urban’ Cypriot Greek from Nicosia (see Terkourafi, 2004, for the description of the urban variety of Cypriot Greek), who was a student at the University of Cambridge. At the time of the study she was 26 years old and had been living in the UK for four years. EE did not report any speech or hearing disorders.

2.1.1.3. Procedure of recording

Written instructions in Cypriot Greek were given to the speaker along with the two test sentences. The speaker was instructed to produce the sentences six times as naturally as possible at a convenient rate of speech, without accelerating.

The recording, which lasted about fifteen minutes, took place in the sound-insulated booth of the Phonetics Laboratory, University of Cambridge. The speaker was recorded using a Sennheiser, model MKH 40 P48, condenser microphone with cardioid characteristics and a Symetrix SX 202 microphone amplifier. The audio signal was recorded to hard disk through the line input of the audio interface of a Silicon Graphics O2+ workstation. The application software used for recording was Silicon Graphics' 'mediarecorder' configured for wav format, 22.05KHz sample rate and 16 bit sample width.

2.1.1.4. Stimuli

The stimuli for the perceptual study were created from the recordings of EE. From the test sentences of Table 1 only the first clause before the pause, i.e. 'I didn't say *pít(t)a*', was used as stimulus; this way the only element that would differentiate the two sentences would be the quantity of the alveolar stop in *pít(t)a*.

Of the two versions of the first clause, the one containing the singleton stop was selected, and, with the use of PRAAT scripts, the 'duration tier' (in PRAAT terminology) of the stop was manipulated to lengthen the closure duration of [t] by increments of 30 ms, in order to produce four stimuli ranging from 30 ms to 120 ms. The duration of the aspiration of each of these stimuli was manipulated again, in order to produce four stimuli of increasing duration of aspiration by steps of 20 ms, thus ranging from 10 ms to 70 ms. In manipulating the aspiration of the stops, the burst (i.e. approximately the first 10 ms of aspiration) was left intact, thus only the rest of the aspiration was manipulated durationally. With the manipulation of CD and Asp, 16 stimuli were created, one for each combination of CD and aspiration steps.

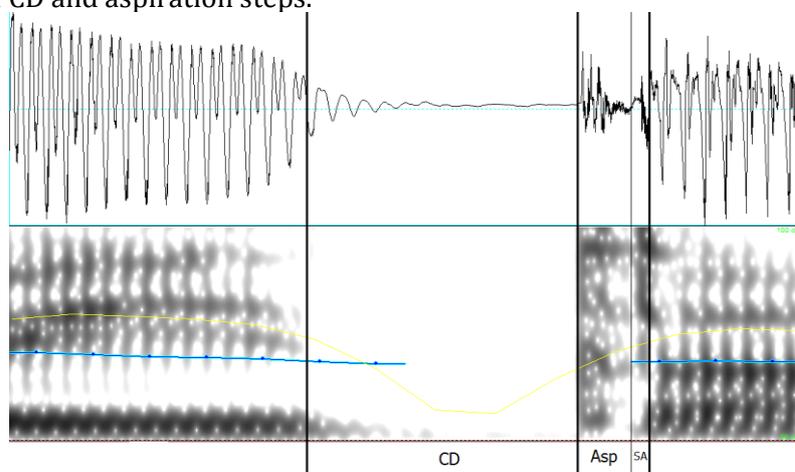


Figure 1: *The original singleton /t/ that served as the basis for the SING-set.*

If the only cues to gemination were to be found in the time domain, then the set of stimuli created would suffice to investigate the perception of those cues. However, in Armosti (2009) it was shown that, acoustically, there were non-durational cues to gemination, such as the intensity of aspiration. Therefore, by selecting the singleton stop to manipulate, there could be some bias caused by the intensity of the singleton stop. In order to account for the non-durational cues found within the consonant, a second set of stimuli was created, this time from the geminate stop,²³ by decreasing the duration of the closure and the aspiration in the reverse of what was done in the case of the first set of

²³ Apart from the spectral differences observable in Figures 1 and 2, the mean aspiration intensity was somewhat lower for the geminate (67 dB) than for the singleton (71 dB).

stimuli (the burst was again left intact). The set of stimuli created from the singleton stop will be hereafter referred to as ‘SING-set’, while the set created from the geminate stop will be referred to as ‘GEM-set’.

A caveat should be mentioned regarding the naturalness of the two sets of stimuli: the spectral characteristics of the aspiration of the singleton (cf. Figure 1) are different from the spectral characteristics of the aspiration of the geminate (cf. Figure 2), thus any lengthening of the aspiration of the singleton would still result in aspiration which would be spectrally different from the aspiration of the geminate. The same holds for shortening the aspiration of the geminate: however short its aspiration may become, it would still carry some spectral characteristics of geminate aspiration, such as differences in high frequency noise. In this sense, an artificial conflict between the duration of the aspiration and its spectral properties is created in some cases. This unnaturalness of those stimuli should be taken into account in analysing the results.

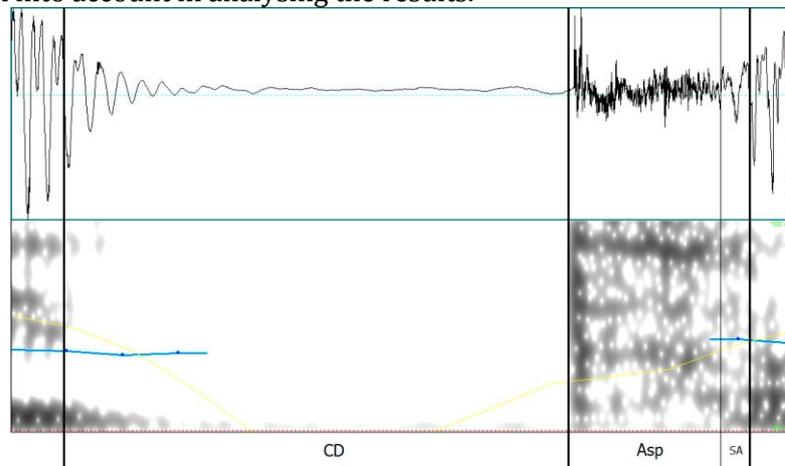


Figure 2: The original geminate /tt/ that served as the basis for the GEM-set.

Regarding the four steps of the manipulated CD and Asp, the reason for selecting the ranges 30 ms – 120 ms for the closure and 10 ms – 70 ms for the aspiration was that those ranges largely coincided with the respective ranges of closure and aspiration duration, as measured for the acoustic study of Armosti (2009) in the case of the word-medial unstressed alveolar stop: 35 ms – 125 ms for CD and 9 ms – 74 ms for aspiration. A graphical representation of these ranges can be seen in Figure 3.

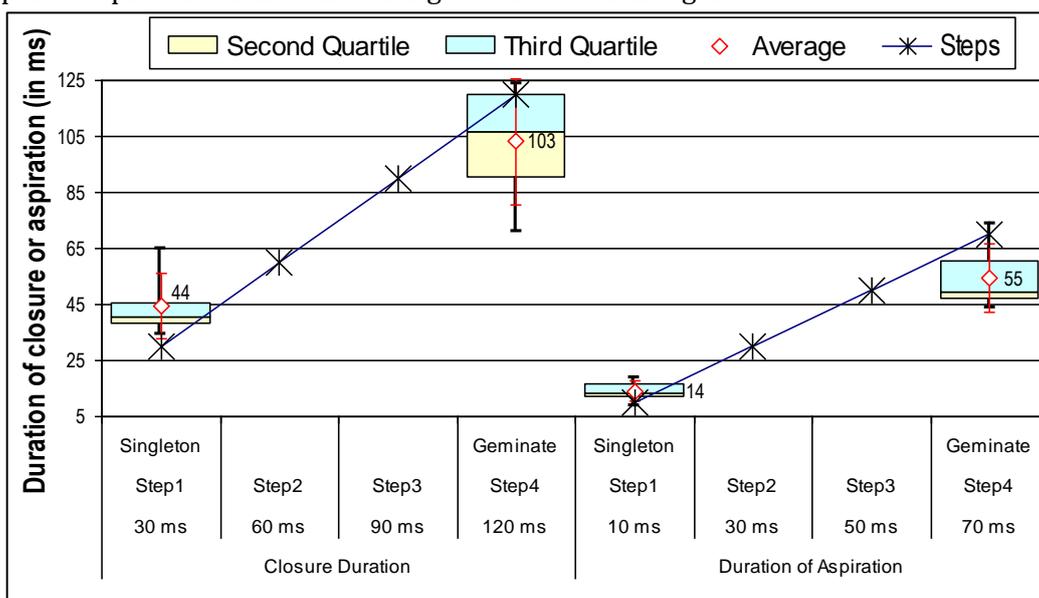


Figure 3: The range of CD and Asp from the acoustic study of Armosti (2009) compared accordingly to the four steps of the stimuli of the present perceptual study.

[Note: The thinner red lines within the box plots indicate a 95% Confidence Interval for the mean; the thicker lines outside the boxes show the first and fourth quartiles]

The 32 stimuli created (2 sets × 4 CD steps × 4 Asp steps) were subjected to one last manipulation: using PRAAT scripts, the manipulated consonant [t^(h)(:)] of each stimulus was extracted and inserted in a single carrier phrase, namely ‘I didn’t say *píta*’, in the place of the singleton alveolar stop of that phrase. The reason for this splicing was to eliminate, in the case of the ‘GEM-set’ of stimuli, any cues to gemination coming from the surrounding vowels (e.g. V1 or V2 duration, pitch, and formants), and, therefore, to concentrate on the properties of the consonant *per se*.

2.1.1.5. Listeners

The subjects of this perceptual study were 14 female and 16 male native speakers of Cypriot Greek. Their age range was from 21 to 34 ($M = 26$, $SD = 2.9$). The vast majority of the subjects were from Nicosia ($N = 25$); three were from Larnaca and two from Limassol. Most of the listeners were regular residents of Cyprus, while a few were students at the University of Cambridge at the time of the experiment. None of the subjects reported any hearing disorders.

2.1.2. Procedure

The experiment was designed using the DMDX Display Software and run on a portable computer. The task was preceded by an introduction, which aimed to familiarise the subjects with the procedure.

2.1.2.1. Presentation of stimuli

During the introductory phase of the experiment, two scenarios were presented in writing to the subjects, one for each test sentence (all the instructions were written in Cypriot Greek). For the ‘*píta*’ case, the scenario was the following:

Maria told Costas to spray the flowers with water. Instead, Costas hosed the flowers forcefully and broke them.

Subsequently, the whole ‘*píta*’ test sentence of Table 1 was heard, as the reply of Maria to Costas: ‘I didn’t say *hose*, I said *spray*’.

A different scenario was presented in writing for the ‘*pítta*’ case:

Maria told Costas to bring her a cake. Instead, Costas brought her a pie.

Following the written scenario, the ‘*pítta*’ test sentence was heard, again as the reply of Maria to Costas: ‘I didn’t say *pie*, I said *cake*’.

After the presentation of the two scenarios, the subjects were informed that for the rest of the experiment they would only listen to the first half of what Maria said, i.e. the stimulus ‘I didn’t say *pít(t)a*’, and would have to select what the second half was accordingly (see the complementary phrases of Table 2).

The purpose of this design was to induce the subjects to concentrate on the meaning and not so much on the phonetic form; by presenting the two complementary phrases of Table 2 (and not the stimuli *per se* in written form) as the two choices, the subjects would never see the crucial geminates written anywhere. Instead, they would hopefully think in terms of the complementary phrase primed by the stimulus.

Table 2: The stimuli and their primed responses.

stimulus	complementary phrase
I didn’t say ‘hose’	I said ‘spray’
I didn’t say ‘pie’	I said ‘cake’

Following the introduction, a small practice session consisting of six randomly pre-selected stimuli was run. The aim of this session was to familiarise the subjects with the

procedure and not with the stimuli; in order to ensure that, this practice session was kept small, and no feedback was provided for the answers the subjects gave.

2.1.2.1. Experimental task

During the actual experiment, the 32 stimuli were randomised in a block, which was presented five times, each time with a different randomisation.

The subjects were asked to identify the stimulus they heard by selecting one of the two complementary phrases (see Table 2) that appeared on the screen immediately after the stimulus was played. Five seconds were provided for the subjects to respond by pressing one of two buttons on the keyboard to indicate their answer; if the subjects did not answer within the five seconds, the program would automatically continue with the next stimulus. After the subjects' response (or after five seconds had elapsed), one second of silence (during which the screen was cleared) followed before the next stimulus was played.

The stimuli were automatically randomised within the five blocks every time the experiment was run. Hence, there were 5 repetitions \times 4 steps of CD \times 4 steps of Asp \times 2 SING/GEM sets = 160 repetitions. The subjects were allowed to have a short break after every block. The total time of the experiment was approximately 15 minutes.

2.2. Analysis

The raw results of the experiment were exported from DMDX in the form of a delimited text file, which was subsequently opened in MS Excel 2003 for processing.

2.2.1. Measurements

Two different measurements were taken for each repetition of the stimuli: (i) the kind of answer the subjects gave, and (ii) their response time. It is hypothesised that the stimuli that would sound more unnatural would require more time for the subjects to respond.

2.2.2. Statistics

The statistical analysis was run in SPSS. The variables and further particulars of the various analyses will be presented before each statistical test in the results section.

3. Results

Two main tests were performed in order to investigate the factors that played a role in the perception of gemination: (i) a Multivariate Analysis of Variance (MANOVA), which tested both the responses and the reaction times of the subjects, and (ii) a logistic regression, in order to determine a descriptive model for the perception of the various cues to germination.

3.1. MANOVA for responses and reaction times

3.1.1. Test variables

'Gemination scores' (see below for its calculation) and reaction times (RT) were the two dependent variables for the MANOVA. The independent variables were: (i) the closure duration, with four levels corresponding to the four steps (30 ms, 60 ms, 90 ms, and 120 ms), (ii) the duration of aspiration, with again four levels for the four steps (10 ms, 30 ms, 50 ms, and 70 ms), and (iii) the 'origin' with two levels, i.e. the SING-set and GEM-set of stimuli.

The 'gemination scores' were initially calculated as the percentage of identification of a stimulus as geminate out of its five repetitions; if a subject did not respond in time for a certain repetition of the stimulus, that repetition was not counted towards calculating the percentage. However, expressing the variable as percentages makes it unsuitable for statistical analysis, as proportional scales are not normally distributed around the mean. Studebaker (1985) proposed a data transformation especially for proportional scales (like the ones found in acoustic and perceptual studies, as he notes), which normalises the data,

and hence makes them suitable for statistical analysis. This method is based on arcsine transformation, with a further linear transformation to make the transformed units numerically close to the original percentages (and thus easier to interpret than mere arcsine units). Studebaker named this method ‘rationalized arcsine transform’, and the units ‘rationalized arcsine units’ (rau).

The second dependent variable, i.e. reaction times, had to undergo a transformation also, as its distribution across the three dependent variables was skewed, therefore not normal. A power transformation was used in order to reduce the skewness of the data, and thus allow it to be used in the statistical analysis.²⁴

3.1.2. General findings

The multivariate tests indicated that all three factors (Asp, CD, and Origin) significantly influenced the way the listeners responded to the stimuli [$F(6, 1856) = 115.324$ for CD, $F(6, 1856) = 144.578$ for Asp, $F(2, 927) = 305.771$ for Origin; $p < .0005$ in all cases]. As shown from their F values, the effect of CD and Asp on the two dependent variables was nearly the same. The univariate results for the two dependent variables will be presented separately.

3.1.3. Gemination scores

The three factors played a significant role on the identification of the stimuli as geminate or singleton, as shown from the results of the univariate tests on the ‘gemination scores’ (see Table 3).

Table 3: *The results for the gemination scores.*

Factor	F value	significance
CD	$F(3, 928) = 366.274$	$p < .0005$
Asp	$F(3, 928) = 526.168$	$p < .0005$
Origin	$F(1, 928) = 584.038$	$p < .0005$
Origin × CD	$F(3, 928) = 23.114$	$p < .0005$
Origin × Asp	$F(3, 928) = 48.975$	$p < .0005$
CD × Asp	$F(9, 928) = 16.861$	$p < .0005$
Origin × CD × Asp	$F(9, 928) = 18.634$	$p < .0005$

It appears that, at the perceptual level, not only the duration of aspiration, but also the duration of the closure plays a significant role in distinguishing between singletons and geminates.

²⁴ As noted in the literature, distributions of response times tend to be L-shaped, i.e. right-skewed (see Bradley, 1975, 1982). Even though such skewed distribution is difficult to be normalised, Box and Cox (1964) proposed a family of power transformations for the normalisation of those distributions: the Box-Cox formula is $y = (x^\lambda - 1)/\lambda$, where x is the original distribution, y the transformed one, and λ the power of the transform (the zero power is taken to be the log x logarithm); the more the original distribution is skewed to the right, the smaller the value of λ must be to obtain a near-normal transformation. Even though Box and Cox provided a sophisticated way to calculate an optimal λ , for the current study the common practice of trial and error (aided by visual inspection of the transformed distributions) was followed until a suitable transformation was found.

Table 4: *Analysed percentages of geminate identification of the stimuli.*

		Duration of aspiration			
		10 ms	30 ms	50 ms	70 ms
CD	30 ms	2%	18%	43%	52%
	60 ms	1%	32%	62%	78%
	90 ms	12%	71%	86%	91%
	120 ms	49%	90%	95%	97%

[Note: The darker the cell of the table, the more the stimulus was identified as geminate.]

Actually, as shown in Table 4, when CD was at its minimum (30 ms), the identification of the stimuli as geminates did not exceed the chance level regardless of the length of their aspiration (only when Asp was at its maximum, i.e. 70 ms, did the gemination score reach the chance level). The same was true for when aspiration was at minimum, i.e. 10 ms.

As shown in Table 3, the origin of the stimuli (i.e. whether the stimuli were created from a singleton or a geminate stop) had a significant impact on the way the two sets of stimuli were perceived, and, moreover, it interacted with the other two factors. This finding suggests that the perceptual weight of the CD and Aspiration cues was different for the two sets of stimuli. Therefore, the two sets should be explored separately.

3.1.3.1. The SING-set of stimuli

An ANOVA test with 'gemination scores' as its dependent variable, and CD and Asp as the two independent variables, was performed for the SING-set of stimuli. Once again, the two factors and their interaction were highly significant, as shown in Table 5.

Table 5: *Results of the gemination scores for the SING-set of stimuli.*

Factor	F value	significance
CD	$F(3, 464) = 253.921$	$p < .0005$
Asp	$F(3, 464) = 127.456$	$p < .0005$
CD × Asp	$F(9, 464) = 12.658$	$p < .0005$

It appears that CD was more important than Asp in the case of the SING-set of stimuli (as indicated by the difference in the *F* values in Table 5). The same conclusion can be drawn from Table 6: for the two smallest values of CD, only one stimulus was identified as geminate above chance level (CD = 60 ms, Asp = 70 ms); the rest of those cases did not exceed 30%, regardless of the length of aspiration.

Table 6: *Analysed percentages of geminate identification of the SING-set of stimuli.*

		Duration of aspiration			
		10 ms	30 ms	50 ms	70 ms
CD	30 ms	3%	5%	5%	17%
	60 ms	1%	12%	27%	59%
	90 ms	9%	44%	71%	83%
	120 ms	40%	81%	91%	93%

Games-Howell post hoc tests showed that the four levels of the CD factor were significantly different from one another ($p < .0005$ in all cases). The same was found for the four levels of the Asp factor, the only exception being the 30 ms and 50 ms steps, for which the difference failed marginally to attain significance ($p = .05$).

3.1.3.2. The GEM-set of stimuli

As was the case with the SING-set, the results of the ANOVA test for the GEM-set of stimuli showed highly significant main effects for the two factors (see Table 7).

Table 7: Results of the gemination scores for the GEM-set of stimuli.

Factor	F value	significance
CD	$F(3, 464) = 118.784$	$p < .0005$
Asp	$F(3, 464) = 492.788$	$p < .0005$
CD × Asp	$F(9, 464) = 24.271$	$p < .0005$

Contrary to the SING-set case, for the GEM-set of stimuli the Asp factor played a greater role than the CD factor in geminate perception, as can be inferred from the *F* values in Table 7. This finding can be observed in Table 8, where only four out of 16 stimuli were not identified as geminates above chance level. Games-Howell post hoc tests for CD showed that there was no statistical difference between the steps 30 ms and 60 ms, and the same was found for the steps 90 ms and 120 ms ($p = .134$ and $p = .052$ respectively). In the case of the Asp factor, only the last two steps (i.e. 50 ms and 70 ms) were not statistically different one from the other ($p = .7$).

Table 8: Analysed percentages of geminate identification of the GEM-set of stimuli.

		Duration of aspiration			
		10 ms	30 ms	50 ms	70 ms
CD	30 ms	1%	31%	81%	88%
	60 ms	1%	53%	96%	97%
	90 ms	15%	99%	100%	100%
	120 ms	57%	100%	100%	100%

3.1.3.3. Comparison of the two sets of stimuli

The observed difference between the two sets of stimuli may be due to non-temporal cues to gemination that reside in the aspiration of the GEM-set (and the absence thereof from the SING-set), as shown in §2.1.2. Figures 5 and 6 below present two stimuli of the same duration of closure and aspiration (30 ms CD + 30 ms Asp) but of different origin (SING-set vs. GEM-set). A mere visual inspection of the two figures reveals spectral differences in the aspiration of the two stimuli. If indeed the spectral quality of the aspiration of the GEM-set of stimuli is characteristic for the aspiration of geminate stops, then the longer the aspiration, the more salient the cues to gemination possibly become. This postulation may serve as an explanation for the observed primacy of the Asp factor over the CD factor in the GEM-set of stimuli, and the reverse in the SING-set. Apart from aspiration *per se*, the spectral quality of the superimposed aspiration (SA)²⁵ was different for the two sets of stimuli, as shown in the Figures 5 and 6. SA was shown to be an important acoustic cue to gemination in Armosti (2009), thus it could play a role in the perception of gemination also.

The two sets of stimuli also differed in the intensity of their aspiration. As shown in Figure 4, the stimuli of the SING-set were of higher intensity than the stimuli of the GEM-set regarding the two smaller steps of Asp duration (10 ms and 30 ms). When the length of aspiration was 50 ms, the stimuli of the two sets were of virtually of the same intensity (approximately 64 dB). For the longest step of Asp duration (70 ms), the intensity of the aspiration of the GEM-set was higher than the intensity of the aspiration of the SING-set.

²⁵ SA may be seen as the overlap of aspiration with the following vowel, indicating breathy voice; for definition of SA, see Armosti (2009), Mikuteit and Reetz (2007), and Clements and Khatiwada (2007).

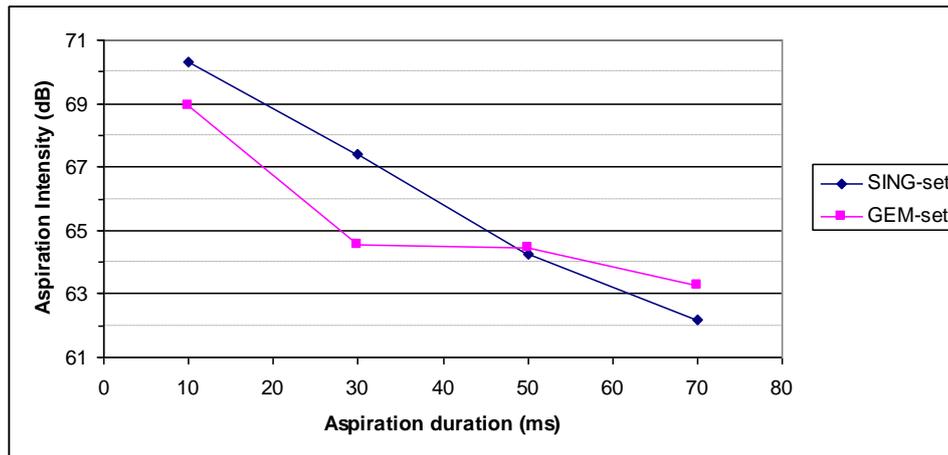


Figure 4: The intensity of the aspiration of the stimuli as a function of the length of aspiration and the stimulus origin.

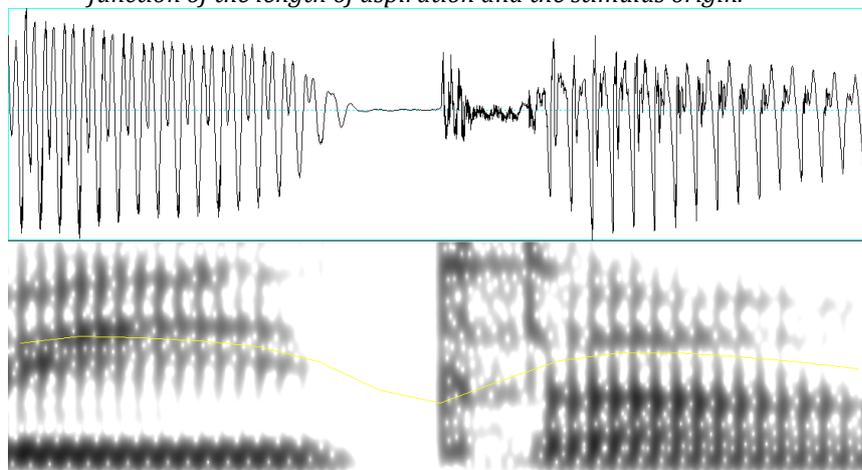


Figure 5: Spectrogram and waveform of the stop in the "30 ms CD + 30 ms Asp" stimulus from the SING-set.

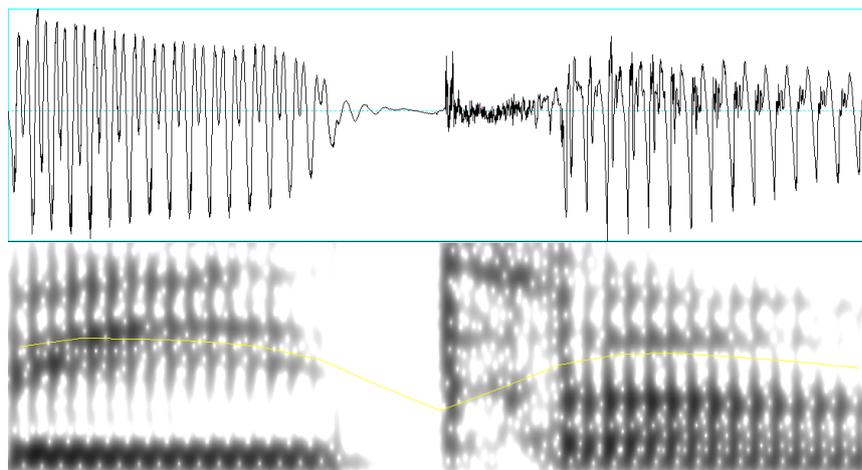


Figure 6: Spectrogram and waveform of the stop in the "30 ms CD + 30 ms Asp" stimulus from the GEM-set.

3.1.2. Reaction times

The univariate tests indicated that the Asp and Origin factors had significant main effects on reaction times, whereas the effects of CD were non-significant.

Table 9: The results for reaction times.

Factor	F value	significance
CD	$F(3, 928) =$	$p = .54$

	0.721	
Asp	$F(3, 928) = 3.192$	$p = .023$
Origin	$F(1, 928) = 29.977$	$p < .0005$
Origin × CD	$F(3, 928) = 14.339$	$p < .0005$
Origin × Asp	$F(3, 928) = 7.568$	$p < .0005$
CD × Asp	$F(9, 928) = 11.156$	$p < .0005$
Origin × CD × Asp	$F(9, 928) = 4.748$	$p < .0005$

Post hoc tests showed that the difference of each level of CD with any other was not significant in any case. For the Asp factor, only the pair 30 ms ~ 70 ms showed a significant difference, with the subjects reacting more slowly for the 30 ms step by 78 ms ($p = .016$).

However, all interactions between the three factors were found significant, and, moreover, planned contrasts revealed that the subjects responded significantly faster for the GEM-set than for the SING-set by 99 ms ($p < .0005$); therefore a separate analysis of the two sets of stimuli is again needed.

3.1.2.1. The SING-set of stimuli

An ANOVA test with RT as its dependent variable, and CD and Asp as the two independent variables was performed for the SING-set of stimuli. CD played a significant role for the speed of the subjects' reaction, whereas Asp did not (see Table 10).

Table 10: RT results for the SING-set of stimuli.

Factor	F value	significance
CD	$F(3, 464) = 5.428$	$p = .001$
Asp	$F(3, 464) = 1.38$	$p = .248$
CD × Asp	$F(9, 464) = 5.858$	$p < .0005$

Post hoc tests showed that there was no difference between the four levels of the Asp factor; in the case of CD, the only significant difference was between the first step (30 ms) and the last two steps ($p = .001$ for the 90 ms step, and $p = .041$ for the 120 ms step). As shown in Figure 7, the subjects were faster at recognising the stimulus “60 ms CD + 10 ms Asp” as singleton (mean RT = 517 ms), and the stimulus “120 ms CD + 50 ms Asp” as geminate (mean RT = 549 ms). For smaller values of CD and Asp, the subjects tended to be faster at perceiving singletons, whereas, in perceiving geminates, they tended to be faster for bigger values of CD and Asp.

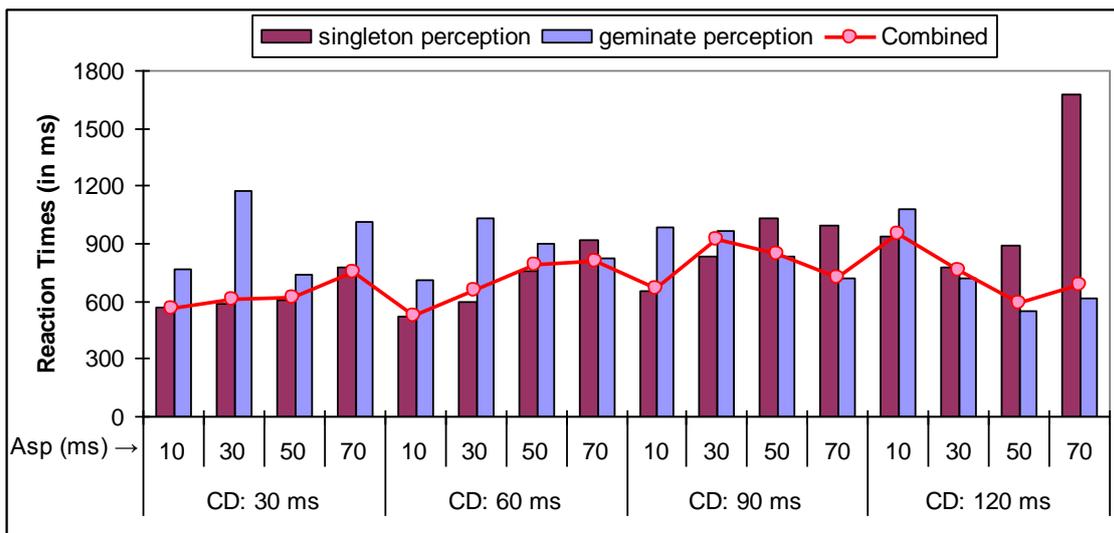


Figure 7: Reaction times for the SING-set of stimuli.

[Note: The bars represent the average RT for the perception of each stimulus as singleton or geminate. The line represents the average RT regardless of the perception of the stimulus.]

3.1.2.2. The GEM-set of stimuli

In the case of the GEM-set of stimuli, both CD and Asp significantly influenced the speed with which the subjects responded; moreover, the effect of the two factors was of the same size, as shown by the *F* values in Table 11.

Table 11: RT results for the GEM-set of stimuli.

Factor	<i>F</i> value	significance
CD	$F(3, 464) = 9.902$	$p < .0005$
Asp	$F(3, 464) = 9.894$	$p < .0005$
CD × Asp	$F(9, 464) = 10.315$	$p < .0005$

Post hoc tests showed that for the GEM-set, the longer the CD was, the faster the subjects replied, with the biggest step (120 ms) having significantly faster responses than the two smaller ones ($p < .0005$ for the 30 ms step, and $p = .041$ for the 60 ms step). The same was found for Asp: the longer the aspiration, the faster the subjects responded, with the biggest step (70 ms) having significantly faster responses than all the rest ($p < .0005$ for the 10 ms and 30 ms step; $p = .047$ for the 50 ms step). Thus the two factors seem to cause the same effect on RT, as shown by both the *F* values of the ANOVA tests, and by the post hoc tests.

As shown in Figure 8, the subjects were faster at recognising the stimulus “120 ms CD + 70 ms Asp” (i.e. the longest stimulus) as geminate (mean RT = 390 ms), and the stimulus “90 ms CD + 30 ms Asp” as singleton (mean RT = 510 ms).

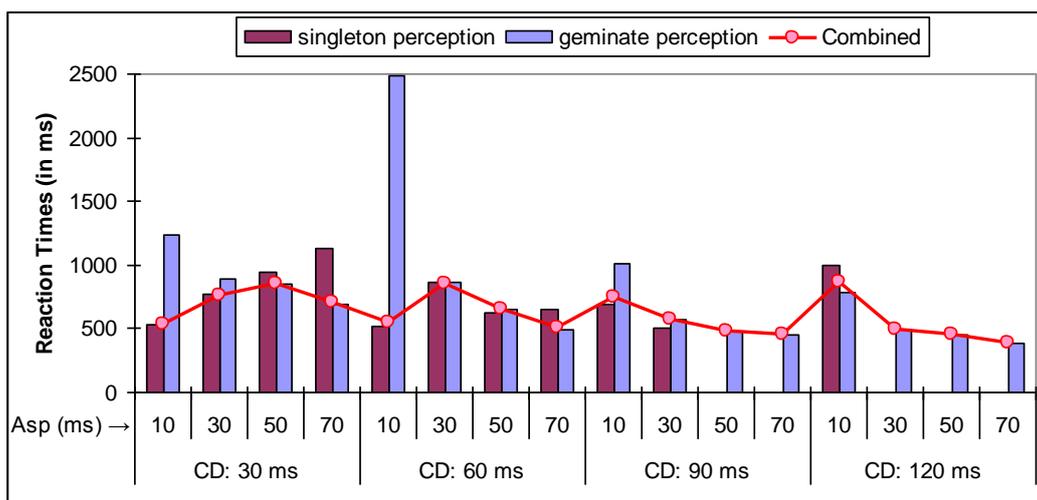


Figure 8: Reaction times for the GEM-set of stimuli.

3.1.2.3. Comparison of the two sets of stimuli

The results regarding RT for the two sets of stimuli showed a correspondence with the respective results for gemination scores (see §3.1.3). For the SING-set, only CD played a significant role for RT, while, for gemination scores, CD had a bigger effect than Asp. Moreover, the stimuli that were more recognised either as singleton or geminate in Table 6 were the ones with the fastest reaction times (see Figure 7).

In §3.1.3.3 it was suggested that the two sets of stimuli differed with regard to a possible presence of non-temporal cues to gemination residing in the aspiration of the GEM-set. The absence of those cues in the SING-set could explain the primacy of the CD factor over the Asp factor observed for the SING-set: since the subjects did not hear those non-temporal cues they expected to find in the aspiration of the stops, they concentrated more on the CD cue, and, therefore, their answers and speed of answering were primarily regulated by the amount of CD they heard.

As for the GEM-set, both factors had a significant role on both the answer given and the speed of answering, even though, for the former, Asp had a greater effect than CD, whereas, for the latter, the two factors had almost the same effect. Accordingly, the stimuli that were faster identified as singleton were of both short CD and aspiration (as shown in Figure 8), while, among them, the ones that were more identified as singletons were the ones of primarily shorter aspiration (see Table 8). Similarly, the stimuli that were faster identified as geminate were of both long CD and aspiration, while, among them, the ones that were more identified as geminates were the ones of primarily longer aspiration.

These findings lend support to the view that the longer the aspiration of the GEM-set, the more salient the non-temporal cues to gemination are, a situation that appears to enhance the aspiration cue relatively to the CD cue. Thus, the subjects focused more on the duration of aspiration (and arguably on the non-temporal cues thereof) and less on CD in identifying the stimuli as singleton or geminate. However, Asp and CD influenced equally the speed of answering, as the longer they were, the faster the subjects answered (see §3.1.4.2), a fact that implies that, even though the aspiration cue was enhanced by the presence of non-temporal cues, CD was still an important cue to gemination. The importance of CD in the identification of the stimuli as singleton or geminate might have been reduced by the presence of those extra cues, but, nevertheless, CD was important enough a cue to provide more confidence (or confusion) to the subjects, hence influencing their speed of answering.

3.2. Modelling the perception of plosive gemination in CyGr

After having identified the factors that play a role in the perception of the stimuli, determining the exact way in which they contribute to that perception was considered to be of interest in order to investigate the relative importance of those cues.

To pursue this investigation, it was necessary to explore the data to identify all possible cues that may influence the perception of a stimulus as a singleton or geminate. The importance of those cues was subsequently evaluated using a logistic regression analysis.

3.2.1. Exploring the data

The aim of the regression analysis was to provide an insight into what acoustic cues the listeners focused on in perceiving an alveolar stop as geminate or singleton. In section 0, three factors, namely CD, Asp, and Origin, were identified as important. Origin was a factor specific to the experimental design of this study: it was meant to distinguish between the two sets of stimuli produced by manipulation of the stop duration according to the source utterance used as the basis of the manipulation. Since the aim of the regression analysis was to infer a model for the perception of alveolar stops based on acoustic cues found in the signal, such an artificial factor as the Origin factor could not be considered for this investigation.

However, it was postulated in section 3.4.1.3 above that the significant contribution of the Origin factor to the perception of the stimuli may be due to cues accumulating in the aspiration portion of the GEM-set. Those cues may be the intensity of aspiration and the length of superimposed aspiration.²⁶ Those two cues were found in the acoustic study of Armosti (2009) to differ significantly between singletons and geminates.

Another cue to gemination found in Armosti (2009) was the total duration of the stop. For the perceptual experiment, a graphical representation of geminate perception with relation to total segment reveals apparent correlations. As shown in Figure 9, the geminate perception of the stimuli of the SING-set increased proportionally to the increase of the total stop duration. In the case of the GEM-set, the increase of geminate perception was again proportional to the increase of total stop duration; the only exceptions were the stimuli with the shortest aspiration (i.e. 10 ms), for which there was a dramatic drop in geminate identification. However, the drop was not random: the identification of stimuli from the GEM-set as geminate in those cases was virtually the same as the identification of stimuli of the same total duration from the SING-set (these cases are indicated with circles in Figure 9). Thus, these drops in the GEM-set can be seen as an exception to the apparent pattern followed by both sets, namely that the total duration of the stop is proportional to its perception as geminate. Therefore, total stop duration was considered to be included in the regression analysis.

²⁶ For the creation of the stimuli, SA was considered to be part of aspiration and not of the following vowel; therefore the length of SA was subject to the overall manipulation of the duration of aspiration.

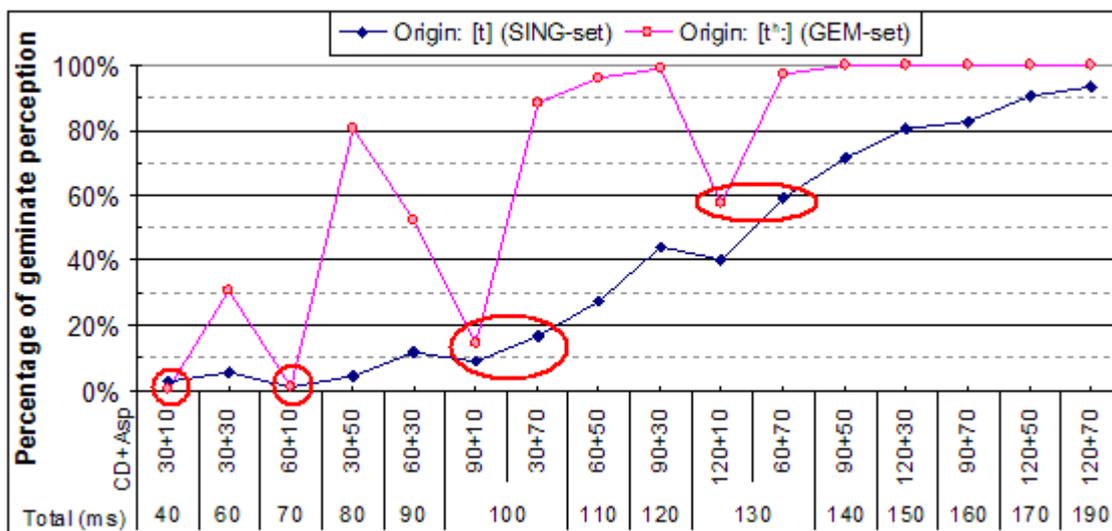


Figure 9: Percentages of geminate identification of the stimuli over total segment duration.

Apart from cues at the segmental level, the acoustic analysis of Armosti (2009) showed that there are some important cues at the supra-segmental level, be they absolute (such as the ‘aspiration plus the following vowel’ sequence, i.e. HV2), or relative (such as the V2:HV2 and V2:CV2 ratios).²⁷ Even though, for the perceptual experiment, V2 was kept the same in all stimuli, the manipulation of the duration of CD and aspiration caused changes in the supra-segmental timing as well. As a result, these changes in HV2, V2:HV2, and V2:CV2 could influence to some degree the perception of the stimuli. Thus, these cues were considered for the regression analysis also.

With the addition of the supra-segmental cues, eight cues in total were regarded as useful for the regression analysis (and were therefore obtained by segmenting and measuring the properties of the stop and V2 of the stimuli): (i) CD, (ii) duration of aspiration (DUR_{Asp}), (iii) intensity of aspiration (INT_{Asp}), (iv) total segment duration (DUR_{Total}), (v) duration of SA, (vi) HV2, (vii) V2:HV2, and (viii) V2:CV2. However, it was expected that some of these variables could be closely connected (i.e. correlated) with one another (such as aspiration with HV2); this situation is called ‘multi-colinearity’, and could produce problems for the regression analysis that would follow.

One way to avoid (or reduce) multi-colinearity is to merge cues that correlate with one another into a single factor. The statistical method to achieve this is called ‘principal component analysis’ (PCA). PCA essentially extracts underlying factors²⁸ from clusters of variables (in this case, from the eight acoustic cues). The analysis extracted two factors, as can be seen from Figure 10: one correlated highly with the total consonant duration and V2:CV2, while the other correlated highly with the duration of aspiration, the aspiration intensity, HV2, and V2:HV2; CD and SA did not correlate with any other variable, hence they were excluded from the two factors.

The clustering of those specific variables into the two factors can be meaningful: the first factor is relevant to the whole duration of the segment, whereas the second factor is closely related to properties of the aspiration. Therefore, the two factors were named ‘Total Segment Factor’ and ‘Aspiration Factor’ respectively. Principal component analysis

²⁷ V2:HV2 is the ratio of V2 to the HV2 sequence; V2:CV2 is the ratio of V2 to the ‘whole stop plus V2’ (i.e. CV2) sequence. These ratios were shown in Armosti (2009) to be acoustic correlates to gemination.

²⁸ Strictly speaking, the outcome of principal component analysis is called ‘component’ and not ‘factor’ (which is the outcome of factor analysis). Nevertheless, the term ‘factor’ is sometimes used generically to include the sense of ‘component’ also.

yielded component score coefficients for each of the two factors, which were imported into their respective formulae (see Equations 1 and 2).

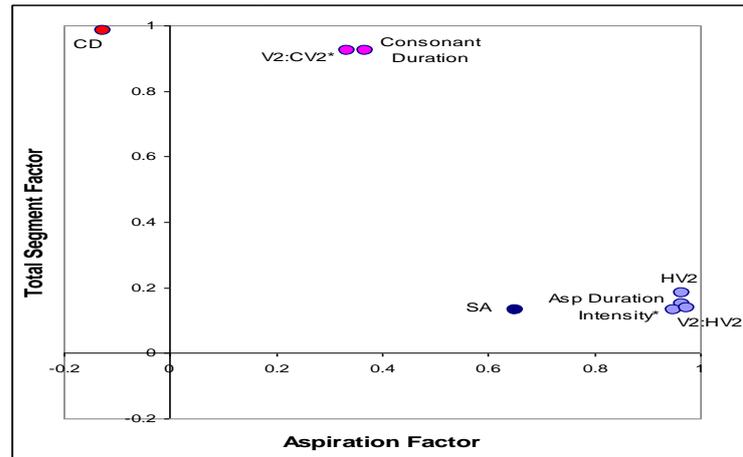


Figure 10: *Component plot of the two extracted factors.*

[Note: The variables indicated with an asterisk (*) show negative correlation with their main factor; however, for the analysis it is the absolute correlation values that are important, therefore those variables were transposed for the clustering to become more apparent.]

$$\text{Aspiration Factor} = 0.256 \text{ DUR}_{\text{Asp}} + 0.005 \text{ DUR}_{\text{Total}} - 0.251 \text{ INT}_{\text{Asp}} + 0.252 \text{ HV2} - 0.259 \text{ V2:HV2} + 0.006 \text{ V2:CV2}$$

Equation 1: *Formula for calculating the 'Aspiration Factor'*

$$\text{Total Segment Factor} = 0 \text{ DUR}_{\text{Asp}} + 0.498 \text{ DUR}_{\text{Total}} + 0.014 \text{ INT}_{\text{Asp}} + 0.019 \text{ HV2} + 0.007 \text{ V2:HV2} - 0.510 \text{ V2:CV2}$$

Equation 2: *Formula for calculating the 'Total Segment Factor'*

With the extraction of the two factors from six variables (out of the eight original variables), and with the two remaining uncorrelated variables (CD and SA), the logistic regression analysis could then be performed on four variables.

3.2.2. Logistic regression

The aim of the logistic regression was to model the exact way in which the factors identified above contribute to the perception of the stimuli as singleton or geminate. The output of the regression analysis was a formula consisting of the acoustic variables (or 'predictors', as they are called in regression analysis) and their weights; that formula can predict the probability of a certain stimulus being perceived as geminate, after the acoustic properties of that stimulus have been entered into the formula.

Thus, the independent variables (i.e. the predictors) entered in the analysis were: (i) CD, (ii) the Aspiration Factor, (iii) the Total Segment Factor, and (iv) SA duration. The method used was block entry regression, in which the predictors are entered one after the other (or in blocks); thus, the way in which each variable contributed to the overall model could be examined.

The analysis showed that all predictors improved the model, except for the Total Segment Factor, which reduced the percentage of correct prediction of the model. A careful examination of the correlations between the four predictors revealed that the Total Segment Factor correlated highly with CD ($r = -.999$), and the Aspiration Factor ($r = -.997$). Correlation between predictors is undesirable in regression analysis, thus the Total Segment Factor had to be removed from the analysis. The remaining predictors were all significant in their contribution to the model, as shown in Table 12.

Table 12: *The results of logistic regression.*

	<i>b</i>	<i>(SE)</i>	Sig	e	95% CI for exp <i>b</i>		
					n.	xp <i>b</i>	Low
Included							
CD	0.035	(0.001)	<.0005	1.036		1.033	1.039
Aspiration Factor	0.11	(0.005)	<.0005	1.116		1.105	1.127
SA	0.197	(0.03)	<.0005	1.218		1.149	1.29
Constant	-5.9	(0.176)	<.0005	0.003			

Note: $R^2 = .345$ (Hosmer & Lemeshow), $.378$ (Cox & Snell), $.506$ (Nagelkerke).
 Model $\chi^2(3) = 2278.384, p < .001$. Percentage of correct prediction: 79.2%.

The percentage of correct prediction of the model was 79.2%. The resulting formula from the regression analysis is shown in Equation 3.

$$P(\text{gem}) = \frac{1}{1 + e^{-(-5.9 + 0.035CD + 0.11AspirationFactor + 0.197SA)}}$$

Equation 3: *Logistic regression formula.*

If the Aspiration Factor in Equation 3 is substituted by its components of Equation 1, then an expanded formula including all eight cues can be derived (see Equation 4).

$$P(\text{gem}) = \frac{1}{1 + e^{-(-5.9 + 0.035CD + 0.028DUR_{Asp} + 0.001DUR_{Total} + 0.027INT_{Asp} + 0.028HV2 - 0.028V2:HV2 + 0.001V2:CV2 + 0.197SA)}}$$

Equation 4: *Expanded regression formula.*

If the acoustic properties of a given stimulus are entered in Equation 4, then the probability of that stimulus being perceived as geminate is generated. For instance, if the properties of the “30 ms CD + 10 ms Asp” stimulus from the SING-set (i.e. $DUR_{Total} = 40$ ms, $INT_{Asp} = 70$ dB, $HV2 = 96$ ms, $V2:HV2 = .9$, $V2:CV2 = .69$, $SA = 4$ ms) are entered in the equation, then the probability of that stimulus being perceived as geminate would be 4.52% (i.e. 95.34% chance to be perceived as singleton). This result is actually very close to the observed result of Table 6 (which was 2.67%).

Testing the predictions of the model against the actual responses of the subjects is a good way to assess the accuracy of the model. However, a model should not only be assessed on the grounds of its true positives (in this case, the success in identifying a geminate), but also of the false positives (i.e. identifying a singleton as a geminate). A method of assessing both aspects of model accuracy is the ‘Receiver Operating Characteristic’ (ROC) analysis. This analysis produces a ROC curve, which is a visual index of the accuracy of the model.

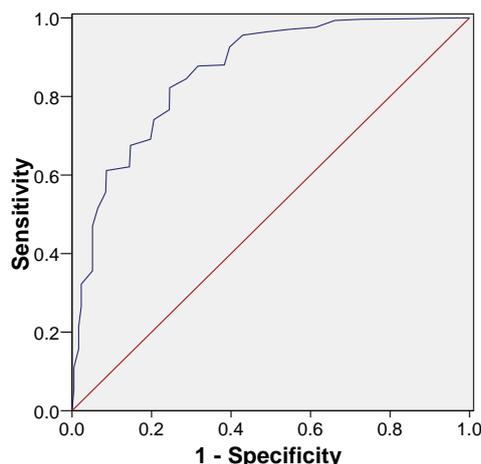


Figure 11: ROC curve for assessing the regression model.

The resultant ROC curve is shown in Figure 11. The term ‘sensitivity’ refers to the accuracy of the model in predicting true positives, i.e. that a geminate is indeed a geminate. The term ‘specificity’ refers to the accuracy in predicting true negatives, i.e. that a singleton is indeed a singleton. For ROC plots, the reverse of true negatives, i.e. false positives (singletons predicted to be geminates), are used; hence the horizontal axis is named ‘1 - specificity’.

The area under the curve represents the probability that the result of Equation 4 for a randomly chosen geminate will exceed the result for a randomly chosen singleton. In other words, the area measures discrimination, that is, the ability of the test to correctly classify singleton and geminate stimuli. The diagonal reference line represents the scenario where the model is not better than guessing, i.e. the discrimination of the two cases is based on chance (the area under the reference line is .5, i.e. 50% chance). The further the ROC curve lies from the diagonal line, the more area it covers, therefore the more accurate the test is. In this study, the area under the curve was .86, which is a good level of accuracy. The asymptotic significance associated with the area statistic was less than .05, which means that using the model is better than guessing.

3.2.2.1. Relative weight of predictors

As mentioned earlier in section 1, previous studies have argued about the relative importance of the CD and aspiration cue to gemination, claiming that one or the other is the primary one. However, arguing about the possible primacy of the one or the other cue should be principally associated with perception rather than production; even though these studies did make inferences regarding the perceptual importance of those cues, their results were in most cases based purely on acoustic data. The present study is the only one that has investigated the perception of those cues through manipulation of their duration, and therefore its claim to reveal the relative importance of those cues has greater validity.

An evaluation of the relative importance of the predictors can be reached by comparing their coefficients in the regression formula, as the one with the greater coefficient contributes more to the model. However, the coefficients are not directly comparable if the predictors do not come from the same underlying distribution. As shown in Figure 3 in section 3.1.2, the values of the duration of the closure and aspiration fall into ranges of different size (the range of aspiration is smaller than the range of CD); therefore, a small increase in aspiration may have a different effect compared to the same increase in CD. Moreover, if these two factors were to be compared with the intensity of aspiration, no comparison could be made, as intensity is measured in different units than duration. A method to avoid these obstacles in comparing different predictors is the standardisation of their distributions; this method ensures that the distributions will have the same mean ($M = 0$) and standard deviation ($SD = 1$).

The predictors were therefore standardised and the regression analysis was run again with the new variables. The results of the analysis are shown in Table 13. The new regression coefficients for CD and the Aspiration Factor were more similar (1.28 and 1.328 respectively) than before (cf. Table 12). These values are also shown in the new regression formula (Equation 5). Their exponential values in Table 13 represent the odds of geminate against singleton perception after a change of one standardised unit of the predictor. Thus, if CD increases by one standardised unit, the odds of perceiving a geminate are 3.6 times higher than perceiving a singleton; when the Aspiration Factor increases by one standardised unit, then the odds are 3.8. This means that the two predictors influence the perception of geminates in almost the same way, with the Aspiration Factor exhibiting marginally more perceptual weight.

Table 13: *The results of the standardised logistic regression.*

	<i>b</i>	<i>(SE)</i>	<i>Sig</i>	<i>e</i>	95% CI for <i>exp b</i>	
					<i>Lower</i>	<i>Upper</i>
Included						
CD	1.28	(.047)	<.0005	3.597	3.279	3.946
Aspiration Factor	1.328	(.059)	<.0005	3.775	3.363	4.237
SA	0.229	(.051)	<.0005	1.257	1.138	1.388
Constant	0.331	(.039)	<.0005	1.393		

Note: $R^2 = .353$ (Hosmer & Lemeshow), $.385$ (Cox & Snell), $.515$ (Nagelkerke).
 Model $\chi^2(3) = 2333.929$, $p < .001$. Percentage of correct prediction: 79.2%.

$$P(\text{gem}) = \frac{1}{1 + e^{-(0.331 + 1.28CD + 1.328\text{AspirationFactor} + 0.229SA)}}$$

Equation 5: *Standardised regression formula.*

The Aspiration Factor, though, is a combination of other variables (such as the duration and intensity of aspiration). The studies that contrasted closure and aspiration did so at a durational level only (they compared CD not with aspiration, but with VOT, which is by nature a durational cue). To allow for a comparison of the perceptual weight of all cues, the regression formula had to be expanded, as shown in Equation 6.

$$P(\text{gem}) = \frac{1}{1 + e^{-(0.031 + 1.28CD + 0.34DUR_{Asp} + 0.007DUR_{Total} - 0.333INT_{Asp} + 0.334HV2 - 0.344V2:HV2 + 0.009V2:CV2 + 0.229SA)}}$$

Equation 6: *Standardised expanded regression formula.*

The regression coefficients and their exponential values are shown in Table 14, which ranks the predictors according to the absolute values of the coefficient *b*. It becomes obvious that, at the durational level, CD has by far more perceptual weight than aspiration: the change in odds for CD is 3.6, whereas for DUR_{Asp} is only 1.4. This means that an increase of CD by one standardised unit induces geminate perception at a much greater degree (around 2.5 times) than the same increase in the duration of aspiration. Actually, the duration of aspiration has virtually the same perceptual weight with V2:HV2, HV2, and INT_{Asp} . Of these four predictors, V2:HV2 is marginally more important, albeit negatively correlated with geminate perception.

Table 14: *The ranking of the predictors according to their perceptual weight.*

Rank	Factor	coefficient	GEM odds	SING odds
		<i>b</i>	Exp <i>b</i>	1 / Exp <i>b</i>
1	CD	1.28	3.597	0.278
2	V2:HV2	-0.344	0.709	1.411

3	DUR _{Asp}	0.34	1.406	0.711
4	HV2	0.334	1.397	0.716
5	INT _{Asp}	-0.333	0.717	1.395
6	SA	0.229	1.257	0.796
7	V2:CV2	0.009	1.009	0.991
8	DUR _{Total}	0.007	1.007	0.993

4. Discussion

The main objective of this paper was to investigate the relative importance of acoustic cues in the perception of alveolar stops as singleton or geminate. This was achieved mainly with the logistic regression analysis, which was supported by the results of the multivariate analysis of variance.

The main finding was that, at a purely durational level, CD was a more important cue to gemination than aspiration (see Table 14). This was mainly shown by the regression analysis when the various constituents of the Aspiration Factor were isolated, so as to discern the perceptual weight of the duration of aspiration regardless of the rest of the cues related to aspiration (such as the intensity of aspiration). In a way, this separation was also achieved in designing the experiment: the stimuli that originated from the singleton stop (i.e. the SING-set of stimuli) were deprived of the non-temporal cues to gemination, and, therefore, only the durational dimension of the aspiration and closure could play a role in the perception of geminates. Indeed the MANOVA test showed that for the SING-set both the duration of the closure and aspiration played a significant role in the answer the subjects gave, with CD being somewhat more important than Asp. Moreover, Asp did not play a role in the speed with which the subjects responded for the SING-set, whereas CD did play a significant role. Therefore, the lack of any non-temporal cues to gemination in the SING-set induced the subjects to rely on CD more than Asp in both the answer they gave and their reaction times.

However, this separation of the duration of aspiration from the rest of the cues that relate to aspiration is only artificial, as, in reality, a geminate stop in Cypriot Greek differs from a singleton not only in terms of the duration of the aspiration and closure, but regarding other acoustic cues, as was shown in Armosti (2009). Those cues were preserved in the stimuli that originated from the geminate stop (i.e. the GEM-set of stimuli), and, indeed, the analysis showed that, even though both CD and Asp were significant in the perception of the stimuli of the GEM-set as geminate or singleton, the effect of Asp was bigger than CD; regarding reaction times, both cues were significant, but had virtually the same effect. When the various aspiration-related cues were treated as a single factor in the regression analysis, similar results were found: both CD and Aspiration were significant predictors in the perception of the stimuli, with Aspiration having marginally more perceptual weight than CD.

A secondary finding was that relative supra-segmental timing is not only *acoustically* an important cue to gemination (as shown in Armosti 2009), but also *perceptually*, as the V2:HV2 ratio was found to be the second most important predictor in geminate perception. However, since V2 remained artificially the same for all stimuli, the importance of this finding may be questioned.

Regardless of the limitations of the present study concerning V2 and the unnaturalness of some stimuli, its results are consistent with the findings of previous (acoustic) studies, namely that not only aspiration, but also CD has a contrastive role in the plosive system in Cypriot Greek (and, therefore, there is a gemination contrast in Cypriot Greek plosives, and not merely an aspiration contrast).

5. Conclusion

This paper provided important findings regarding the debate about which cue should be considered primary for the perception of stops in Cypriot Greek. In particular, it showed that both CD and aspiration are important *perceptual* cues to gemination, with aspiration exhibiting marginally more perceptual weight than CD. Interestingly, the duration of aspiration *per se* (i.e. without considering the other cues related to aspiration, such as its intensity) is substantially of less importance than the duration of the closure. However, since aspiration influences perception not only by its duration, but also by the other cues related to it, the finding regarding aspiration as a combination of cues (i.e. that it is slightly a better predictor for gemination than CD) should be of more relevance to the debate.

Regardless of which of the two elements of the stop is more important for the perception of gemination, what this study has demonstrated about CD is that it cannot be denied that CD is an important perceptual cue to gemination in Cypriot Greek—nearly as important as aspiration. This finding can serve as a further indication towards analysing the plosive system of Cypriot Greek as one contrasting (unaspirated) singletons with (aspirated) geminates, rather than merely unaspirated with aspirated plosives. If the latter had been the case (as Davy and Panayotou, 2004, and Charalambopoulos, 1982, argued), then CD would not have exhibited such a perceptual weight.

Consequently, the perception of plosive gemination in Cypriot Greek is partly determined by CD, i.e. the universal main cue to gemination; however, in Cypriot Greek aspiration serves as a marginally more important cue than CD.

References:

- Armosti S. (2009). *The phonetics of plosive and affricate gemination in Cypriot Greek*. Unpublished PhD dissertation, University of Cambridge.
- Arvaniti A. & Tserdanelis G. (2000). On the phonetics of geminates: evidence from Cypriot Greek. *Proceedings of the 6th International Conference on Spoken Language Processing* (Beijing, 6–20 October 2000): 559–562.
- Boersma P. & Weenink D. (2007). Praat: doing phonetics by computer (Version 4.5.17) [Computer program]. Retrieved March 2007 from <http://www.praat.org/>.
- Botinis A., Christofi M., Themistocleous C., & Kyprianou A. (2004). Duration correlates in Cypriot Greek. *Proceedings of the 17th Swedish Phonetics Conference FONETIK 2004* (Stockholm, 26–28 May 2004), 144–147.
- Box G. E. & Cox D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 24(2): 406–424.
- Bradley J. V. (1975). The optimal-pessimal paradox. *Human Factors* 17: 321–327.
- Bradley J. V. (1982). The insidious L-shaped distribution. *Bulletin of the Psychonomic Society* 20: 85–88.
- Charalambopoulos A. (1982). Συμφωνικός «διπλασιασμός» και δασύτητα στην Κυπριακή [Consonant “doubling” and aspiration in Cypriot]. *Studies in Greek Linguistics* 3: 237–255.
- Christodoulou C. (2007). Phonetic geminates in Cypriot Greek: the case of voiceless plosives. *Proceedings of the 8th Annual Conference of the International Speech Communication Association, Interspeech* (Antwerp, 27–31 August 2007): 1014–1017.
- Clements G. N. & Khatiwada R. (2007). Phonetic realization of contrastively aspirated affricates in Nepali. In J. Trouvain & W. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (Saarbrücken, 6–10 August 2007): 629–632.
- Davy J. & Panayotou A. (2004). Phonological constraint on the phonetics of Cypriot Greek: Does Cypriot Greek have geminate stops? *Proceedings of the 6th International Conference of Greek Linguistics* (Rethymno, 18–21 September 2003). Rethymno: Linguistics Lab: 1–13.
- Forster K. I. & Forster J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35: 116–124.
- Mikuteit S. & Reetz H. (2007). Caught in the ACT: The timing of aspiration and voicing in East Bengali. *Language & Speech* 50(2): 249–279.

The perception of plosive gemination in Cypriot Greek

- Muller J. S. (2001). *The phonology and phonetics of word-initial geminates*. Unpublished PhD dissertation, Ohio State University.
- Studebaker G. A. (1985). A “Rationalized” Arcsine Transform. *Journal of Speech and Hearing Research* **28**: 455-462.

The phonology and phonetics of glides in North-Western Greek dialects

MARY BALTAZANI
University of Ioannina
mbaltaz@cc.uoi.gr

NINA TOPINTZI
Aristotle University of Thessaloniki
topintzi@enl.auth.gr

1. Introduction

Apart from reports on high vowel deletion and mid vowel raising in Northern Greek dialects (Chatzidakis 1905, Papadopoulos 1927, Newton 1972, Browning 1991, Kondosopoulos 2000, Trudgill 2003) there is hardly any description of the phonetic quality of vowels²⁹—and even less so of glides—surfacing in these dialects. Standard Modern Greek (SMG) has been reported to have only one glide, [j], in its inventory³⁰ (Mirambel 1959, Householder 1964, Newton 1972, Joseph & Philippaki-Warburton 1987), while in Northern Greek, in addition to the high front glide, there have been anecdotal, impressionistic reports about the existence of a high back glide, [w] (Phavis 1951, Newton 1972).

Crosslinguistically, segments labeled as *glides* have variable phonological and phonetic patterning, in that they display both consonantal and vocalic characteristics, sometimes becoming part of consonant clusters, while at other times forming diphthongs with vowels (see Nevins & Chitoran 2008 and references therein). The most common glides across languages are [j] and [w], which are thought to be closely related to [i] and [u] respectively.

Turning to the Greek phonological literature, there is no consensus on the phonological status of glides in SMG. Some scholars claim that [j] is an allophone of /i/ (Newton 1961, Kazazis 1968, Warburton 1976, Malavakis 1984, Nikolopoulos 1985), others argue for the existence of two separate phonemes /i/ and /j/ (Mirambel 1959, Koutsoudas 1962, Householder *et al.* 1964, Setatos 1974, Nyman 1981), while a third proposal puts forth the idea of an underlying archi-phoneme /I/ which is underspecified for the feature [consonantal] and which relies on the ‘Maximal syllabification principle’ to account for the surface realization of the segment sometimes as a vowel and sometimes as a glide (Deligiorgi 1987, Malikouti-Drachman & Drachman 1990). Clarifying the phonological status of the [j] glide in SMG is beyond the purposes of our study, but we will add another piece to this puzzle by describing a different type of glide in North-Western Greek (NWG), not attested in SMG, which has a different behaviour (section 2.1).

So far, no phonetic investigation of the dialectal glides has taken place to our knowledge, which is an essential step before any further analysis is undertaken. This is one of the aims of this paper, together with their phonological investigation as well as a comparison between NWG and SMG, which will hopefully promote discussion on Greek glides in general. To sum up, our aims in this paper are (a) to distinguish among different types of NWG glides and establish the phonetic environments where they appear; (b) determine whether this phenomenon in NWG is categorical or variable; (c) tentatively seek the reasons behind its different realizations.

²⁹ See, however, Trudgill (2009) for a recent analysis of the vowel system of the greek dialect spoken in Sfakia.

³⁰ This description has been questioned in phonetic studies (Malavakis 1984; Arvaniti 1999, 2007; Nicolaidis 2003) which show that what is phonologically described as /j/ is realized phonetically as a voiced fricative [j̥].

2. Phonetic description of glides in NWG

The phonetic characteristics of Northern Greek glides, as already mentioned, have by and large received no attention till now. The only notable exception is a brief *impressionistic* description found in Newton (1972) who reports of a high front glide [j] and a high back glide [w]: “All dialects have the high front glide [j] ... and many have a high back glide [w].” (1972: 11). However, Newton hasn't got much to say about glides in specific dialects or the phonetic environments they occur in besides a remark of Phavis (1951) who observes glide formation before stressed mid vowels reporting: “...a pronunciation [wó] for [ó] in Kozani and other parts of Macedonia.” (Newton 1972: 29).

The current work offers a first analysis of these glides. The material we base our analysis on comes from a corpus of spontaneous and semi-spontaneous speech recordings of 12 speakers from the area of Western Macedonia (Kozani) and Epirus (Ioannina and Arta). The speakers were all in the 50-60 year old range and they reported, through conversation, on everyday matters, childhood memories, war memories etc. for about 60 minutes with the interviewer, with minimal interruption. For this paper, 5-6 minutes from 4 speakers were analyzed, in which we counted 125 tokens containing glides.

The next section supplies the results of this investigation. Starting from general observations about NWG glides, we then move on to their acoustic analysis (§2.1.1). §2.1.2 deals with the distributional properties of glides introducing us to the topic of the next section (§3), namely, a discussion on their phonological status.

2.1. The results

Our first finding is in accordance with previous impressionistic reports that NWG dialects present two glides, namely the palatal [j] and the labio-velar [w]. The second finding is more surprising; in particular, we offer evidence suggesting that NWG distinguishes between two types of glides: the first type of glides is common between NWG and SMG, appearing in exactly the same positions in both dialects (henceforth COMMON); the second type on the other hand is idiosyncratic to NWG and does not appear in comparable positions in SMG (termed here NWG-ONLY). In (1) we give examples of words containing the COMMON type of glides and in (2) the NWG-ONLY glides and compare them to their corresponding SMG words.

(1) Words with COMMON glides in NWG

NWG	SMG	Gloss
piðjá	peðjá	'children'
tsimbjéndan	tsibjótan	'was enamored'
ðjo	ðjo	'two'
çérja	çérja	'hands'

(2) Words with NWG-ONLY glides and their correspondents in SMG

NWG	SMG	Gloss
kwókaʎu	kókalo	'bone'
mwóʎis	mólis	'just before, as soon as'
pwósa	pósa	'how many'
mjésa	mésa	'inside'
ksjéru	kséro	'I know'
patjéra	patéra	'the father'
funájɟ	fonázi	's/he shouts'

Note that the transcription of glides in (1) and (2) is different. We transcribe the glides in (1) as fricative palatals [ʎ], while in (2) as approximant palatals [j]. Contrary to what has been reported in Nevins & Chitoran (2008), phonetic studies of SMG note that the SMG palatal glide surfaces as a voiced palatal fricative (Malavakis 1984; Arvaniti 1999, 2007;

Nicolaidis 2003) and our data show the same realization for COMMON glides as well (see Figure 1 in section 2.1.1 below). On the other hand, NWG-ONLY glides do not show any trace of frication (see Figure 2). We discuss these differences in section 2.1.1

We found far more NWG-ONLY glides than COMMON ones in our data; 111 and 14 tokens respectively. Before discussing the differences between the two types of glides in terms of distribution, we present their acoustic realization in NWG.

2.1.1. Acoustics of NWG glides

As is well-known, the acoustic structure of glides, or semi-vowels, corresponds to that of vowels. For instance, Ladefoged & Maddieson (1996: 323) point out that “...within each language the semi-vowels differ from the corresponding vowels in that they are produced with narrower constrictions...”, hence, the formant structure of [j] and [w] roughly corresponds to that of [i] and [u] respectively. In addition, the narrower constriction of /j/ in turn often leads to palatalization and/or affrication of a preceding consonant (Hall & Hamann 2006; Hall et al. 2006) – which is the predominant realization in COMMON cases – as was reported in 2.1 above (see (1)). Figure 1 gives a representative example of a word containing a COMMON type glide in the word *tsibjjendan* ‘got enamored’ realized [tsimbjendan] in NWG. The frication of the glide is evident after the voiced stop [b].

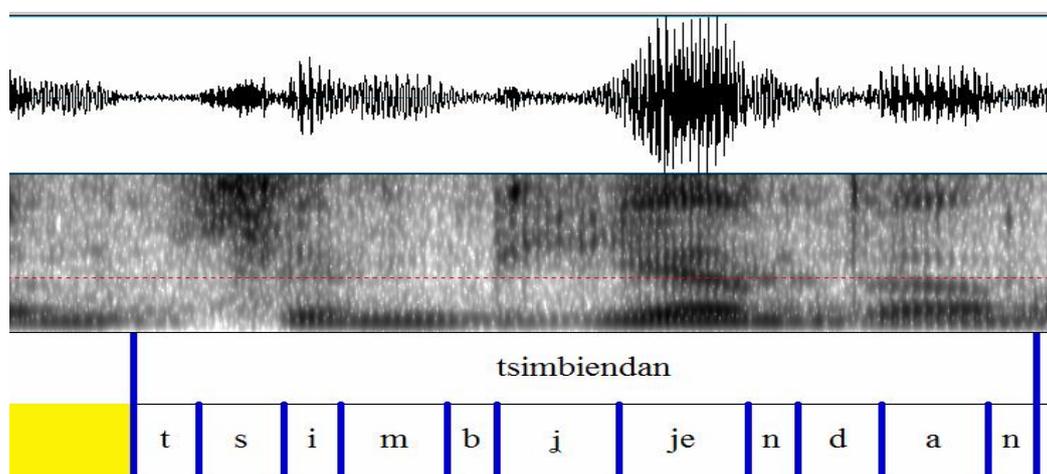


Fig. 1: The NWG word [tsimbjendan] *tsimbjotan* ‘got enamored’ shows the fricative portion of the glide immediately after [b].

Figure 2 shows an example of an NWG-ONLY glide in the word *enas* ‘one’ realized [jenas] in NWG. There is clearly no frication in the glide realization here and this difference is consistent between COMMON glides and NWG-ONLY ones, as is also shown in Figures 3, 4 and 6 below.

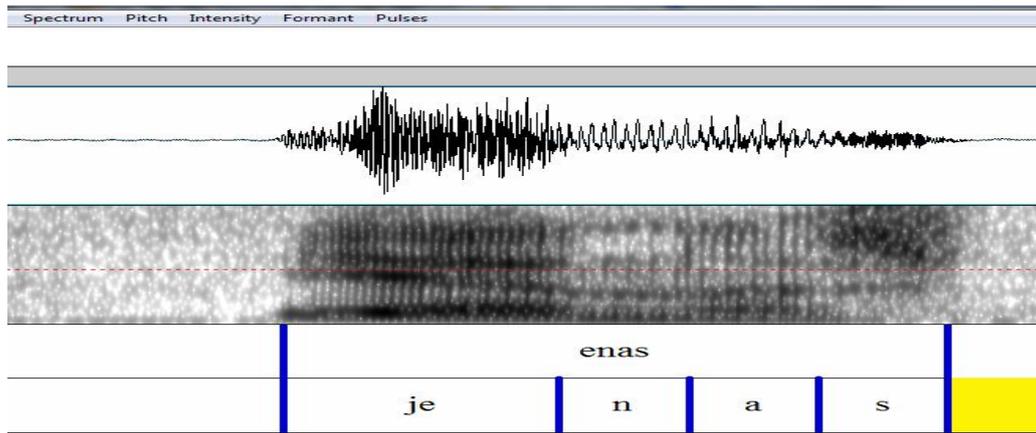
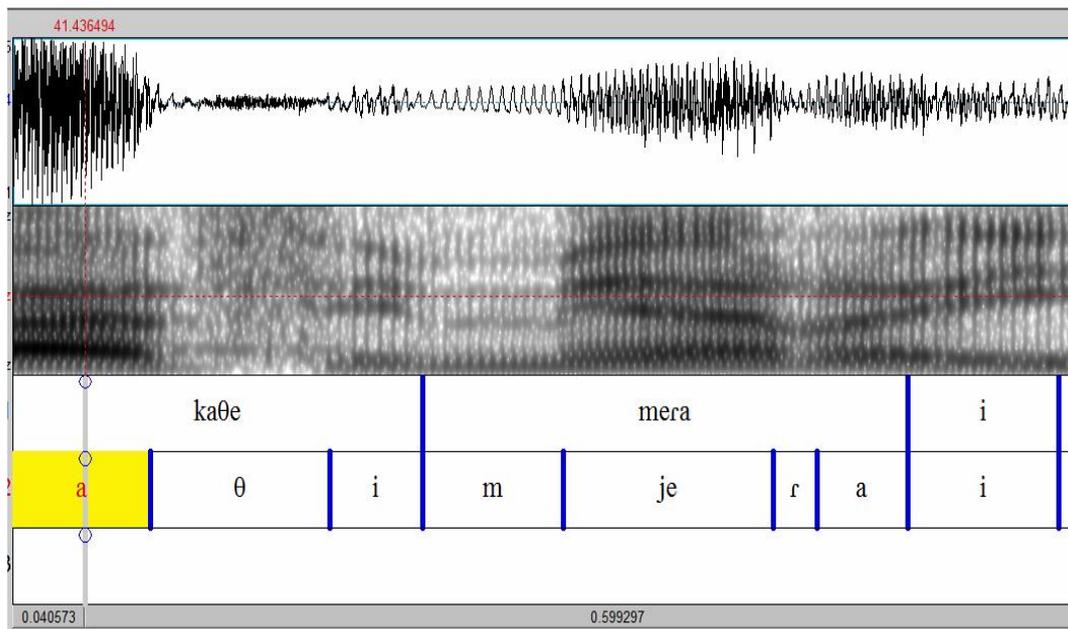


Fig. 2: The NWG word [jenas] *enas* ‘one’ shows that there is no frication in NWG-ONLY glides in word-initial environments.

Arguably, the frication part could be missing from the token of Figure 2 because the glide is not postconsonantal. However, our data show that there is no frication in NWG-ONLY glides in any environment, as is clearly evident in Figure 3 which shows two representative tokens of post-consonantal NWG-ONLY glide: in the words *mera* ‘day’ realized [mjera] in NWG (top panel) and *patera* ‘father’ realized as [patjera] (bottom panel). À propos of the example [mjera] we should note another difference between NWG and SMG: it is very common for a [mj] cluster in SMG to be realized with an epenthetic [ɲ], that is, [mɲj]; that is not the case for NWG-ONLY glides, which as the example in Figure 3 shows has no such epenthetic segment.



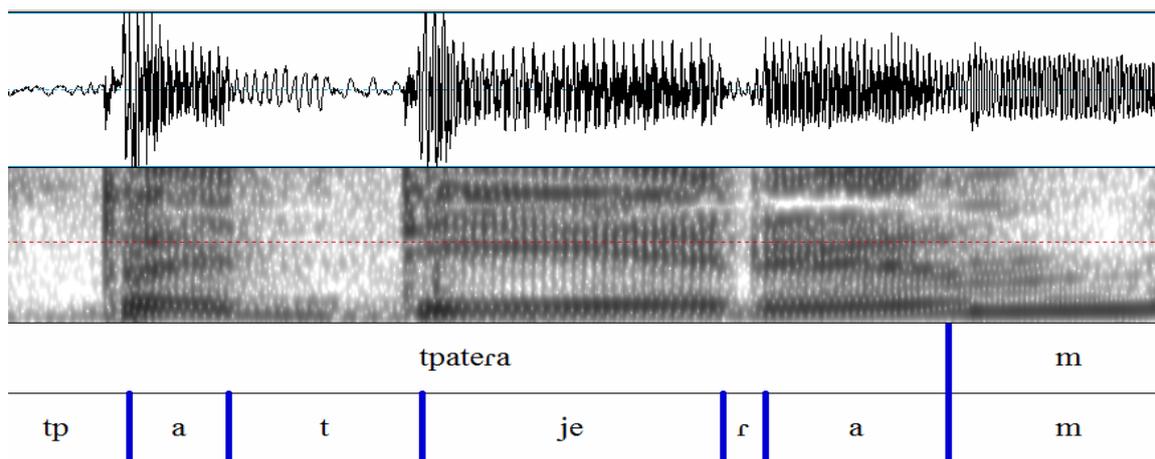


Fig. 3: The NWG words [mjera] *mera* 'day' (top) and [patjera] *patera* 'father' (bottom) showcase realization of NWG-only glides in a postconsonantal/prevocalic environment. No frication is evident.

This salient difference between the two types of glides can be aerodynamically attributed to the high velocity of the airflow produced at the release of a stop which is higher with greater constriction degrees of the following vocoid (Ohala 1983, Nevins & Chitoran 2008). In other words, the phonetic realization of our data suggest that NWG-ONLY glides do not show the frication part because they are more vowel-like (have smaller constriction)³¹ than the COMMON glides which are more consonant like (greater constriction).

One further difference in the realization of the two types of glide is regulated by stress: NWG-ONLY glides appear only in stressed syllables, while there is no such restriction for COMMON glides. Figure 4 gives an excellent example of the role of stress in NWG-glides. The speaker self-corrects, changing the position of stress in the word *katevenan* 'they went down'. First he pronounces it [kati'vjen(an)] with penultimate stress and the second time [ka'tjevinan] with antepenultimate stress. This change in stress position brings about the change in the position of glide insertion, as well.

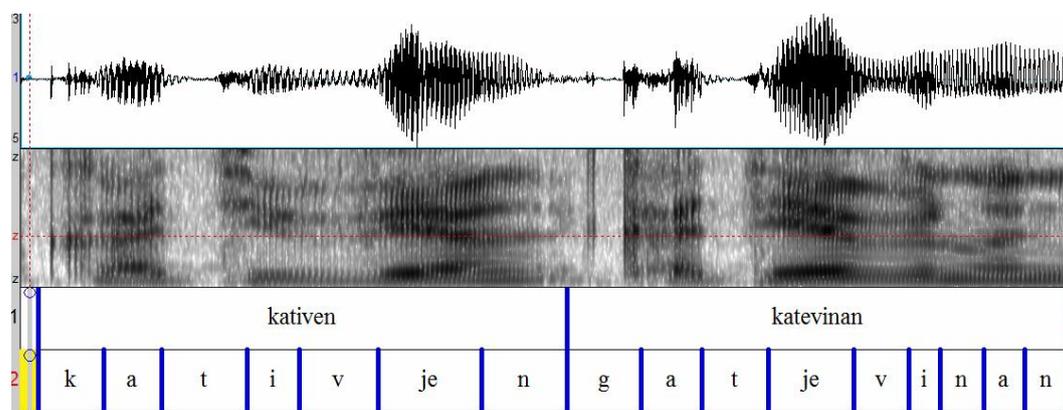


Fig. 4: The role of stress in NWG gliding. On the left, the word *katevenan* 'they went down' is realized [kativjen(an)] with stress and [je] in penultimate position; on the right it is realized [katjevinan] with stress and [je] in antepenultimate position.

³¹ This opens up the possibility that NWG glides function as diphthongs. A similar process appears in Romance. In the Romance languages, the original Latin short vowels /e/ and /o/ have generally become diphthongs, [je] and [wo], when stressed, e.g. Latin *petra* 'stone' and *focu* 'fire' evolved in Spanish as [pjédra] and [fwégo] respectively (Chitoran and Hualde 2007: 46). Perhaps the fact that NWG dialects also have this phenomenon, but SMG does not, has to do with the contact of N. Greece with such languages through the Balkans. Nonetheless, the fact that this process exists in Romance does not explain why it exists, or why it started in the first place. Since we became aware of this possibility at the final stages of writing this paper, we will explore this alternative in future work.

Figure 5 gives an example of an NWG-ONLY [w] glide in the word [fwotu] ‘Fotu’s (name)’. Formant movements are shown to highlight the similarity between the height of F2 at the beginning of [wo] in the first syllable and at the steady state of [u] in the second syllable. Note how F2 rises for the position of [o] near the middle of the [wo] syllable.

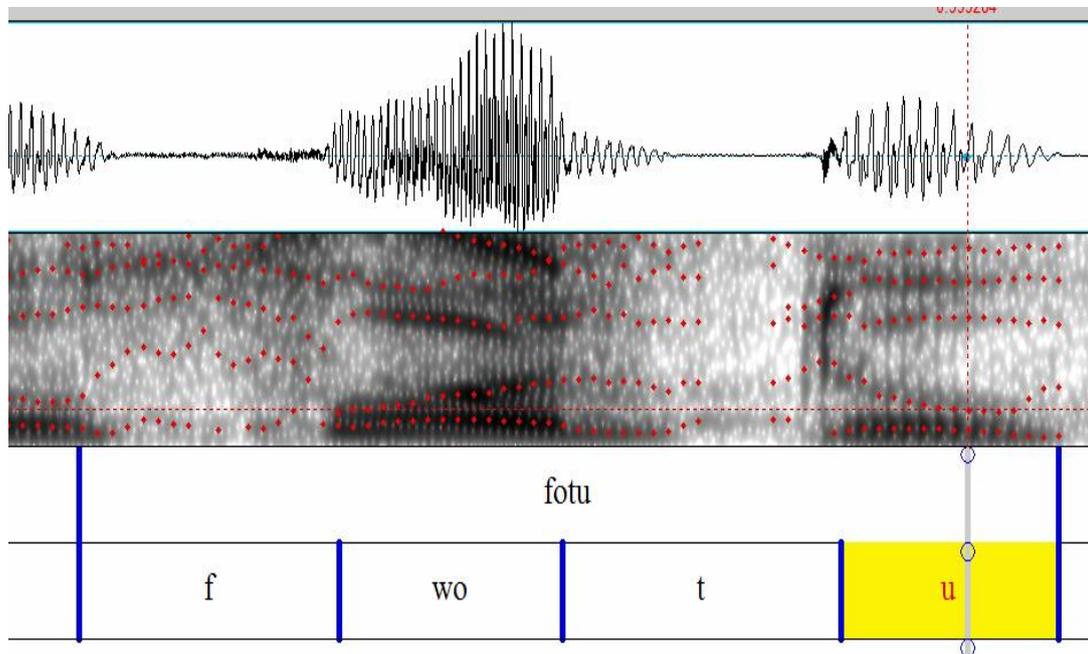


Fig. 5: Example of [wo] in the word [fwotu] ‘Fotu’s (name)’.

Since the glides [j] and [w] have similar formant values to [i] and [u], we expect the formants in sequences [je] and [wo] to show movement from the high vowel values to those of the mid. We measured formant movements of 10 words each for the [je] and [wo] from the NWG-ONLY category. Figure 6 shows the average measurements of F1 (bottom) and F2 (top) taken $\frac{1}{4}$ into the vowel [e] then at the $\frac{1}{2}$ and $\frac{3}{4}$ points over the 10 tokens measured of words with [je]. Movement of F1 and F2 from the values typical for [i] to the values typical for [e] is evident which we interpret as the presence of an onglide to the vowel.

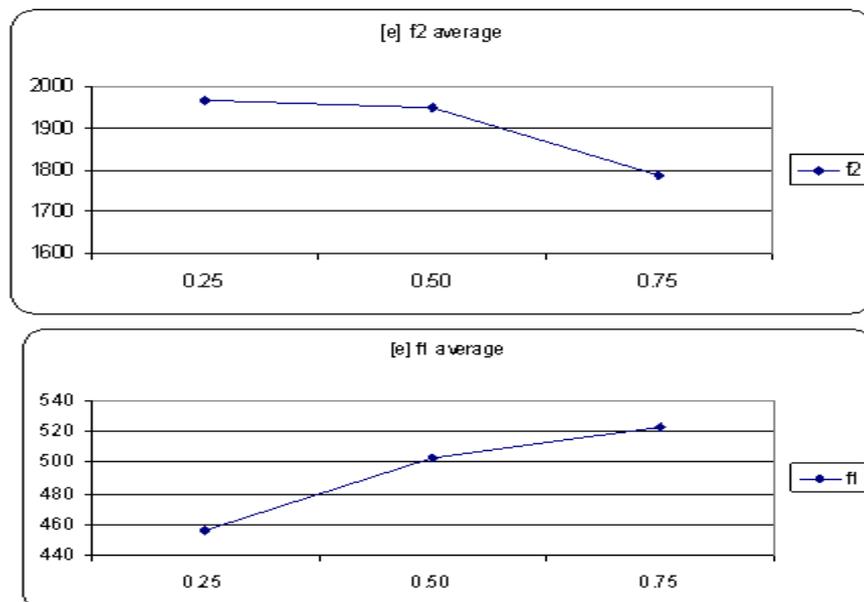


Fig. 6: Movement of F1 (bottom) and F2 (top) from the values typical for [i] to the values typical for [e] in [je] sequences (average from 10 tokens).

Figure 7 shows measurements for [o] taken $\frac{1}{4}$ into the vowel, then at the $\frac{1}{2}$ and $\frac{3}{4}$ points over the 10 tokens measured of words with [wo]. Movement of F1 and F2 from the values typical for [u] to the values typical for [o] is evident (averages over 10 tokens), which we interpret as the presence of an onglide to the vowel.

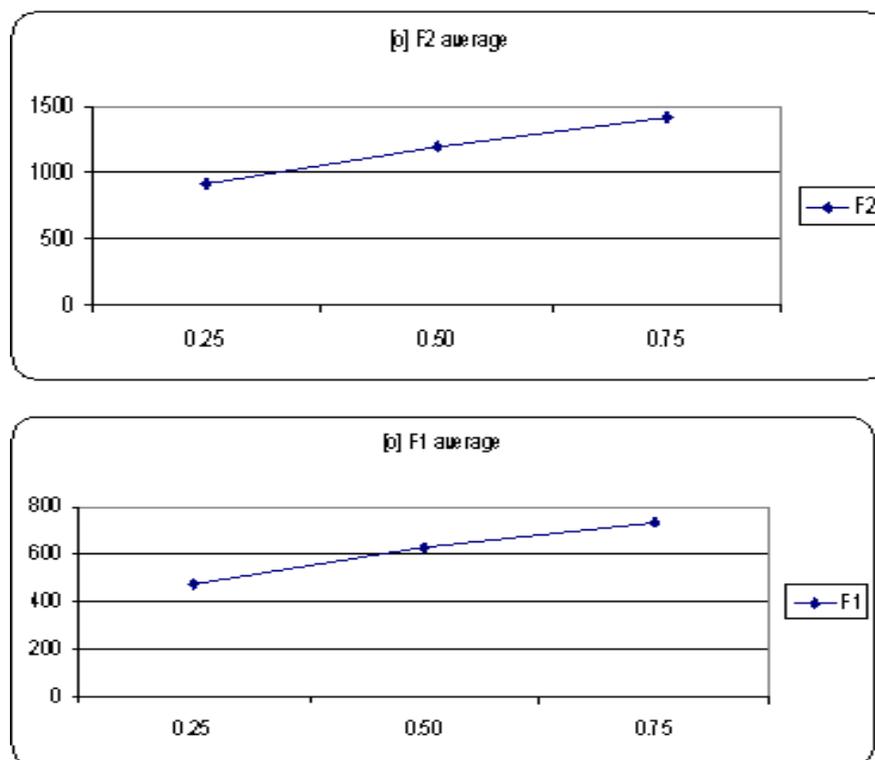


Fig. 7: Movement of F1 (bottom) and F2 (top) from the values typical for [u] to the values typical for [o] in [wo] sequences (average from 10 tokens).

In sum, examination of the acoustic properties of glides in North Western Greek showed that there are two different types of glide in this variety which differ in their acoustic realization: the type which is similar to the SMG glides—and which we called COMMON—is characterized by frication, while the type of glide which is only attested in North Western Greek—the NWG-ONLY glides—is realized without any frication. The former glides are arguably realized with greater constriction and this suggests they are more consonant-like, while the latter with less constriction or more vowel-like. The next section examines the distribution of NWG-ONLY glides both with respect to the COMMON glides as well as with each other, i.e. a comparison between *j* and *w*, something we have not yet discussed.

2.1.2. Distribution

Starting with a distribution comparison of the COMMON and NWG-ONLY glides, we observe that the former category only comprises one glide, i.e. [j], whereas the latter contains both [j] and [w], their distribution being regulated by the following vowel; the mid-front vowel is preceded by [j], while the mid-back one is preceded by [w], cross-linguistically a very common distribution. This distribution holds for the overwhelming

majority of cases³². In *prevocalic* environments [j] and [w] only appear before [i] and [o] respectively. Such a restriction does not hold for the COMMON glide which may be followed by any vowel.

More specifically, in the set of 111 NWG-ONLY tokens, both [j] and [w] emerge with approximately the same frequency: we found 46 [je] tokens (41%) and 45 [wo] tokens (40%). Less frequently, glides appeared postvocally. In particular, we found 12 [aj] tokens (10.8%) only before palatals and 8 tokens (7%) of the type [oj], [uj] and [ej], while only 1 token of [ow]. Finally, NWG-ONLY glides arise much more often as onglides than offglides. In the latter case, they basically appear before palatal consonants only. No similar limitation seems to be pertinent to COMMON glides.

A second important difference relates to the role of stress. The COMMON glide may or may not be found within a stressed syllable, but the NWG-ONLY glide necessarily occurs within a stressed one (cf. (2) above).

Our data also reveal a third difference between the two types of glide concerning their obligatoriness: Words with NWG-ONLY glides display variable realizations, some with and some without the glide; for example, we found instances of the same speaker pronouncing *patera* ‘father’ both as [patera] and [patjera]. On the other hand, words containing COMMON glides are never realized without one.

Table 1 summarizes all the preceding remarks on the differences between COMMON and NWG-ONLY glides.

Table 1: *Distribution of COMMON and NWG-ONLY glides.*

	COMMON	NWG-ONLY
[j]	✓	✓
[w]	✗	✓
Licensed by stress	✗	✓
Obligatoriness	✓	✗ (optional)
Combination with any V	✓	Mainly [je] & [wo]
Onglide position: (j+V, w+V)	✓	✓
Offglide position: V+j	✓	Restricted (before palatals)
V+w		Very rare (1 token in 45)

Table 1 above seems to imply that the NWG-ONLY [j] and [w] pattern in the same way. Although it is true that they share the property of both appearing in stressed syllables, they are different in other respects. More specifically, [j] appears word-initially and word-medially with almost the same frequency, while [w] appears mostly word-medially—we found only 5 word-initial tokens. When prevocalic, [j] comes after any type of consonantal articulation (except velars), but [w] mostly follows labial and velar consonants. The basic observations are summarized in Table 2 below.

Table 2: *Differences in the distribution of NWG-ONLY [j] and [w] in NWG.*

		[j]	[w]
Word-initial		16	5
After	labials	9	16
	interdentals	5	1
	alveolars	6	4
	palatals	10	0

³² There are some cases, approximately 18% of our nwg-only tokens, where [j] appears with other vowels, but crucially in all of these cases it is an off-glide, appearing *after* a vowel (mostly [a]) and before a palatal sibilant [ʃ] or [ʒ]. Due to the scarcity of offglides in our data, the following discussion mainly focuses on onglides.

3. The phonology of glides

3.1. Brief overview of previous studies on Greek glides

Early accounts of glides viewed underlying vowels as the only source for surface glides (e.g. Kaye and Lowenstamm 1984, Steriade 1984, Levin 1985, Rosenthal 1994). In such approaches however, the difference in the consonantal vs. vocalic behaviour of glides was not clearly evident. To capture the duality of glides as consonants or vowels, Clements and Hume (1995) instead assigned particular constituenthood within the syllable or feature structure, so that glides could bear place features under the C- or V-place-nodes. Much more recently, Levi (2008) has attempted to capture differences in glide behaviour in a rather direct approach. In particular, she differentiates between underlying and derived glides. The former refer to 'real' phonemic glides that pattern with consonants, whereas the latter refer to underlying vowels that surface as glides, but pattern with vowels.

This distinction finds equivalents in SMG where a contrast between underlying and derived glides seems extant.

(3) SMG phonologically

Underlying:	/mjalo/	[mjaló]	'mind'	
Derived:	/mati/	[máti]	'eye' but	/mati+a/ [mátja] 'eyes'

Specifically for Greek now, a number of proposals have been put forward to account for glides. The three main approaches are listed in (4) and outlined below (cf. Rytting 2005 for details).

(4) Proposals about Greek glides

- i) Allophonic (e.g. Kazazis 1968, Warburton 1976)
- ii) Phonemic (e.g. Setatos 1974, Nyman 1981)
- iii) Underspecification (e.g. Deligiorgi 1987, Malikouti-Drachman & Drachman 1990)

The former follows the tradition (see Kaye & Lowenstamm 1984 above and others) which claims that glides necessarily come from underlying vowels. Thus, depending on the environment, a high vowel may surface as a vowel or as a corresponding glide, i.e. /i/ → [i] ~ [j]. This allophonic account however misses cases whereby a lexical contrast between vowels and glides arises. Consider for instance the word *áðeia*. For many speakers this is pronounced as [áðia] when it means 'permission' and as [áðja] when it means 'empty-PL-NEUT.'. Minimal pairs of this kind motivate the phonemic account which maps /i/ to [i] and /j/ to [j]. The phonemic proposal is not without problems either, since it fails to capture the cases of derived glides that the allophonic approach so easily accounts for. Lastly, the underspecification account attempts to capture the vowel vs. glide contrast simultaneously with the allophonic relationship by claiming that there is a just a single phoneme /i/ without the need for /j/. The twist required here is that /i/ can either be specified as [-cons], in which case it is systematically interpreted as [i] phonetically or it can be left unspecified for [cons], in which case it can alternate between [i] or [j] depending on the syllabic position.

Despite any advantages each of these accounts has, it is quite clear that each fails to capture a number of facts related to the [i]-[j] alternations, a phenomenon that is usually attributed to socio-linguistic factors or the demotic-katharevousa distinction. Bearing in mind that our purpose here is to describe glides in NWG as adequately as possible – given our presently limited corpus of data – and to remark on their innovations when compared to those of the standard dialect, we currently refrain from reaching any theoretical conclusion on SMG glides and focus instead on certain aspects of the dialectic

glides. For this reason, we do not take a stand as to whether these glides are underlying or derived and will continue using the theory-neutral terms COMMON and NWG-ONLY.

3.2. The NWG-ONLY glides

Argumentation supporting the difference between COMMON and NWG-ONLY glides has been presented in §2.1.1 and §2.1.2. This now brings us to the question; if NWG-ONLY glides are distinct from the COMMON ones, then how can we theoretically analyze them? This is the topic of the next sub-sections.

3.2.1. NWG-ONLY glides function epenthetically

Recall that prevocalic NWG-ONLY glides appear either after another consonant (...CjV...) as in [mjéra] or word-initially (#je or #wo) as in [jétsi] 'so, this way' or [wótan] 'when'. Let us first consider the latter instance which seems more straightforward to account for, since it appears to be driven by the need to satisfy ONSET by means of onset epenthesis. At first sight, no similar justification seems to be available for the /...CV.../ → [...CjV...] change, given that an onset is present already. However, we argue that the epenthesis of a glide in NWG, whether to offer a (new) onset or to form a complex onset is driven by the need to create smoother transitions from and to the syllable nucleus.

This idea is inspired by work by Uffmann (2007) who observes that glottal stops are usually epenthetic word- or foot-initially, whereas glides are usually epenthetic intervocalically. This differentiation on the nature of the epenthetic consonant relates to sonority considerations, since different epenthetic consonants may enhance or reduce the contrast of the preceding/following segment. Given that vowels are prominent segments, the best epenthetic segment in a V_V context is a glide due to its high sonority. Uffmann (2007: 458) thus proposes that “glides are inserted to minimise the contrast to the following or preceding vowel”.

Glides in NWG presumably take on this role intervocalically, but, as we presently claim, also prevocalically in #_V, C_V and post-vocalically in V_C contexts³³. To see why, consider Sonority Sequencing (Clements 1990), whereby sonority must sharply rise from the onset to the nucleus and then gradually lower towards the coda. For this reason, ideal singleton onsets are the ones of the lowest sonority such as the stops *p*, *t*, *k*. On the other hand, ideal singleton codas are the ones whose sonority is lower than a vowel, but still not too low. When we add complex margins to the equation, things get slightly modified. The generated strings will consist of [C₁C₂V] for a complex onset and [VC₂C₁] for a complex coda. Davis and Baertsch (2008) observe that the preferable sonority profile of C₁ and C₂ cross-linguistically is the same across the corresponding positions, namely low sonority for C₁ and high for C₂. This proves quite insightful, when we consider the NWG data. C₂ in complex margins is ideally filled by a high sonority segment, a role that is undoubtedly best fulfilled by an epenthetic glide. We can thus claim that the glide is inserted to achieve the preferable sonority profile, thus accounting for the C_V and V_C environments.

But this does not answer the question of why a glide should be epenthesized in the first place. In the #_V context a low sonority singleton onset would offer the ideal rising sonority slope towards the nucleus, whereas in the ...C_V... context, glide insertion seems redundant, as there is already a good sonority profile available. The answer to both questions comes from a single proposal. In particular, we claim that in NWG more important than a simply good sonority profile is to have smooth transitions from an onset to the peak and from the peak to a coda, whenever possible. Uffmann's proposal about the minimisation of contrast offered by glides in relation to vowels now comes in handy. Epenthetic glides in C₂ position or as singleton onsets serve this function in the best way

³³ In the light of Uffmann's (2007) observations, such claim might seem surprising in the #_V context, but it is actually not, if one takes into consideration the lack of [ʔ] in Greek. As for the preference of using [j] over low-sonority consonants in singleton onset-position, an explanation is offered a bit later in the text.

possible and are thus preferred, even at the expense of a more complex syllable structure in the case of ...C_V... or ...V_C...

3.2.2. Some complexities

Naturally at this point, one may wonder: if NWG-ONLY glides are truly epenthetic, then why don't they appear in front of any vowel? Our answer will be that these glides behave epenthetically, but may only surface under assimilatory conditions. To unravel what this means, consider the context where each of the epenthetic glides emerges. In particular, we find [je] but not *[ji] and [wo] but not *[wu]. The prohibition against high glides and vowels presumably indicates that the high glide acts as a separate root node/segment – hence is epenthetic – that cannot co-occur with a high vowel due to an OCP restriction such as **[+high] [+high]*³⁴. Treating the glide as a separate root node on the other hand fails to explain why it is [j] and not [w] that accompanies the front vowel [e] and vice versa for the back vowel [o]. Moreover, it provides no account as to why the central-back [a] is not preceded by the dialect-only [w]. These points however can be answered, if we assume that the NWG-ONLY glide is actually the product of assimilation to the following vowel in terms of the features [-low]&[α back]. Given that [a] is [+low], then it falls out that it will be not preceded by any glide. At the same time, the feature specification of the mid vowels [e] and [o] in terms of backness will regulate the corresponding glides. A similar interaction between glide epenthesis and assimilation is observed in Chamicuro (de Lacy 2006: 106, 129-130) where the inserted [w] glide takes on its specification by the [+back] or [dorsal] feature of the vowel /a/ that systematically precedes it.

The proposal about epenthetic glides is not unprecedented. For instance, in Brazilian Portuguese, Albano (1999) claims that “epenthetic [j] should be regarded as distinct from ‘true’ [j]” based on phonetic evidence that suggest the former glide is “probably the result of a gradient process that can, in this case, be attributed to gesture overlap”. We have also provided phonetic evidence that indicates a distinction between COMMON and NWG-ONLY glides. While in many cases, phonetically distinct glides also contrast phonologically, cf. Sundanese (Levi 2008), such mapping is not always one-to-one. For instance, Levi explains that in Karuk and Pulaar two phonologically contrastive glides receive the same phonetic realization (a many-to-one phonology-phonetics mapping), whereas in Argentinian Spanish a single phonological glide exhibits different phonetic realisations depending on the environment (a one-to-many phonology-phonetics mapping).

For this reason, Levi (2008) suggests that an exploration of the phonological behaviour of glides, with respect to e.g. the syllable, is a more reliable means so as to classify them in different types³⁵. While we currently lack sufficient data to be able to determine the constituent structure of the NWG-ONLY glides, there are numerous other observations which seem to phonologically distinguish them from the COMMON glides. Some relate to distributional properties mentioned already in (§2.1.2) and will not be repeated here. Additional ones are listed in (5) and briefly explained next. Note that (5) refers only to the phenomena in NWG. We do not make claims about SMG.

(5) *Phonological differences between COMMON and NWG-ONLY glides in the dialect*

<u>The glide:</u>	COMMON	NWG-ONLY
-------------------	--------	----------

³⁴ The tendency against *[ji] and *[wu] in languages such as Ignaciano Moxo is alternatively attributed by Ohala and Kawasaki (1984: 122-3) to the fact that these sequences create minimal modulations in amplitude, periodicity and spectrum.

³⁵ However, even when researchers agree on the phonological status of glides, their position within the syllable is often debatable. For example, Yip (2003: 782) observes that: “...Harris (1983) for Spanish and Bao (2000) for Fuzhou locate them in the Rime; Pike and Pike (1947) for Mazateco and Bao (1990) for Mandarin locate them in the Onset; and Clements (1986) for Luganda and Duanmu (1990) for Mandarin consider them secondary articulations on the onset consonant”.

Undergoes fortition to fricative after obstruents	✓	✗
May be preceded by epenthetic nasal	✓	✗
Is the product of assimilation to the following V	✗	✓

A phonological process that often applies to input glides is that of fortition to fricatives after obstruents, thus [pápja] becomes [pápça] 'duck', [ráfja] → [ráfça] 'shelves', [tétjos] → [tétços] 'such-NOM-MASC', etc (for a somewhat similar process, see Nevins and Chitoran 2008 on Cypriot Greek). No similar fortition is applicable to NWG-ONLY glides in words such as [patjéra] 'father' or [kaθjénas] 'everyone'. In a similar vein, an epenthetic nasal may develop before [m] and the COMMON glide [j], as in [mɲjá] 'one-FEM-SMG', but not if the glide is epenthetic, thus *[mɲjéra] 'day' or *[mɲjéxr] 'until'. Lastly, as explained before, the epenthetic NWG-ONLY glide is the product of assimilation to the features [-low] & [α back] of the following vowel, whereas the COMMON glide seems to be present as such in the underlying representation (see also §3.1 for discussion).

These differences therefore suggest that the distinction between COMMON and NWG-ONLY glides is not only phonetic, but also phonological. Notably, some of the phonological facts are corroborated by the phonetic findings. In particular, we showed in §2.1.1 that COMMON glides bear a frication part that renders them more consonant-like as opposed to the more vowel-like NWG-ONLY glides, which lack this feature. It thus seems no accident that only the former undergo fortition to fricatives – since they are consonant-like – in contrast to the latter ones.

4. Conclusion

In this paper, we have investigated the behaviour of glides in the North-Western variety of Greek. Several new findings have emerged. First, we established acoustically, for the first time to our knowledge, the existence of a high back glide [w] in addition to the high front glide [j]. Secondly, we have shown through phonetic evidence that there are two distinct types of glide in this variety, one that we termed COMMON and another that we termed NWG-ONLY. We showed that these two types differ in three respects: (a) their phonetic realization, in that the former type is realized mostly as a fricative while the latter as an approximant, (b) their distribution, in that the former can occur in any type of syllable, whereas the latter is only found within stressed syllables and (c) their obligatoriness, in that the former is obligatory while the latter is not. Finally, we offered a preliminary account of the phonological structure of the NWG-ONLY onglide and argued that its function is epenthetic but subject to assimilatory conditions as well as to the OCP. This explains why it is found before certain vowels only. We also compared its phonological behaviour with that of the COMMON glide and identified certain differences between them. These findings have thus led us to the claim that the two types of glides are distinct both in terms of their phonetics as well as their phonology.

References

- Albano E. (1999). 'A gestural solution for some glide epenthesis problems', *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, San Francisco, 1785-1788.
- Arvaniti A. (1999). 'Illustrations of the IPA: Modern Greek', *Journal of the International Phonetic Association* 19: 167-172.
- Arvaniti A. (2007). 'Greek Phonetics: The State of the Art', *Journal of Greek Linguistics* 8: 97-208.
- Bao Z.-M. & Fan-Qie (1990). 'Languages and reduplication', *Linguistic Inquiry* 21.
- Bao Z.-M. (2000). 'Syllabic constituency and sub-syllabic processes', *Journal of East Asian Linguistics* 9: 287-313.
- Browning R. (1991). *Medieval and Modern Greek [H ελληνική γλώσσα μεσαιωνική και νέα]*, Athens: Papadima Publications [Greek edition].
- Chatzidakis G. (1905). *Μεσαιωνικά και Νέα Ελληνικά Α'*, P.D. Sakellarios, Athens.
- Chitoran I. & Hualde J.I. (2007). 'From hiatus to diphthong: the evolution of vowel sequences in Romance', *Phonology* 24: 37-75.

- Clements G. N. (1986). 'Compensatory lengthening and consonant gemination in Luganda', In: *Studies in Compensatory Lengthening*, Wetzels, L., Sezer, E. (eds.), Foris, Dordrecht, 37-78.
- Clements G. N. (1990). 'The role of the sonority cycle in core syllabification', In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, J. Kingston and M. E. Beckman (eds.), Cambridge University Press, Cambridge, 283-333.
- Clements G. N. & Hume E. (1995). 'The internal organization of speech sounds', In J. Goldsmith (ed.) *The Handbook of Phonological Theory*, 245-306. Cambridge, MA: Blackwell.
- Davis S. & Baertsch K. (2008). 'On the Relationship between Codas and Onset Clusters', Paper presented at the CUNY Conference on the Syllable, January 17-19, 2008.
- de Lacy P (2006). *Markedness: Reduction and Preservation in Phonology*, Cambridge: Cambridge University Press.
- Deligiorgi I. (1987). 'On rules of syllabification and syllabic restructuring', *Proceedings of Chicago Linguistic Society* 23/2: 76-90, Chicago: Department of Linguistics, University of Chicago.
- Duanmu S. (1990). *A Formal Study of Syllable, Tone, Stress and Domain in Chinese languages*. MIT: PhD Dissertation.
- Joseph B. D. & Philippaki-Warbuton I. (1987). *Modern Greek*, London: Croom Helm.
- Hall T. & Hamann S. (2006). 'Towards a typology of stop assibilation', *Linguistics* 44.
- Hall T., Hamann S. & Zygis M. (2006). 'The phonetic motivation for phonological stop assibilation'. *Journal of the International Phonetics Association* 36.1: 59-81.
- Harris J. (1983). *Syllable Structure and Stress in Spanish*, MIT Press, Cambridge, Mass.
- Householder F.W. (1964). 'Three Dreams of Modern Greek Phonology', In R. Austerlitz (ed.) *Papers in Memory of George C. Papageotes* (issued as Supplement to *Word* 20/3: 17-27).
- Householder F.W., Kazazis K. & Koutsoudas A. (1964), *Reference Grammar of Literary Dhimotiki*. Bloomington, IN: Indiana University Press.
- Kaye J. & Lowenstamm J. (1984). 'De la syllabicité'. In F. Dell, J. Hirst & J.-R. Vergnaud (eds.) *Forme sonore du langage*, 123-160. Paris: Herman.
- Kazazis K. (1968). 'Sunday Greek', *Papers from the Fourth Regional Meeting, Chicago Linguistics Society* (CLS 4), Chicago: Department of Linguistics, University of Chicago, 130-140.
- Kontosopoulos N. (2000). *Διάλεκτοι και Ιδιώματα της Νέας Ελληνικής*, 3η έκδοση, Αθήνα: Εκδόσεις Γρηγόρη.
- Koutsoudas A. (1962). *Verb Morphology of Modern Greek*, Bloomington: Mouton.
- Ladefoged P. & Maddieson I. (1996). *The sounds of the world's languages*, Oxford; Cambridge, MA: Blackwell.
- Levi S. (2008). Phonemic vs. derived glides, *Lingua* 118: 1956-1978.
- Levin J. (1985). *A Metrical Theory of Syllabicity*, MIT: PhD Dissertation.
- Malavakis Th. (1984). 'Φωνηεντικές Συνέχειες: Διφθογοποίηση, Ουρανικοποίηση και Φωνηματική Κατάταξή τους', *Studies in Greek Linguistics* 4: 1-16.
- Malikouti-Drachman A. & Drachman G. (1990). Φωνολογική κυβέρνηση και προβολή: Αφομοιώσεις, ανομοιώσεις. *Working Papers in Greek Grammar*, University of Salzburg, 1-20.
- Mirambel A. (1959). *La Langue Grecque Moderne: Description et Analyse*, Paris: Librairie Klincksieck.
- Nevins A. & Chitoran I. (2008). Phonological representations and the variable patterning of glides. *Lingua* 118: 1979-1997.
- Newton B. (1961). Rephonemization of Modern Greek. *Lingua* 10: 275-284.
- Newton B. (1972). *The Generative Interpretation of Dialect: A Study of Modern Greek Phonology*. Cambridge: Cambridge University Press.
- Nicolaidis K. (2003). Μια Ηλεκτροπαλατογραφική Μελέτη των Ουρανικών Συμφώνων της Ελληνικής. In D. Theophanopoulou-Kontou, C. Lascaratou, M. Sifianou, M. Georgiamentis & V. Spyropoulos (eds.) *Σύγχρονες Τάσεις στην Ελληνική Γλωσσολογία*, 108-127. Αθήνα: εκδόσεις Πατάκης.
- Nikolopoulos G. (1985). Re-examination of stress and diphthongization in Modern Greek [Επανεξέταση του τόνου και της συνίζησης στα Νέα Ελληνικά]. *Studies in Greek Linguistics* 5: 21-40.
- Nyman M. (1981). Paradigms and transderivational constraints: Stress and yod in Modern Greek. *Journal of Linguistics* 17: 231-246.
- Ohala J. (1983). The origin of sound patterns in vocal tract constraints. In: *The Production of Speech*, 189-216. Springer.
- Ohala J. & Kawasaki H. (1984). Prosodic phonology and phonetics, *Phonology Yearbook* 1: 113-127.
- Papadopoulos A. (1927). *Γραμματική των Βορείων Ιδιωμάτων της Νέας Ελληνικής*, Εν Αθήναις: Π.Δ. Σακελλάριος.

- Phavis V. (1951). Ο δυναμικός τονισμός της βορείου Ελληνικής και τα αποτελέσματα αυτού, *Αθηνά* 55: 3-18.
- Pike K. & Pike E. (1947). Immediate constituents of Mazateco Syllables. *I.J.A.L* 13: 78-91.
- Rosenthal S. (1994). *Vowel/Glide Alternation in a Theory of Constraint Interaction*, PhD Dissertation, University of Massachusetts, Amherst. Published 1997, New York and London: Garland Publishing.
- Rytting A. C. (2005). An iota of difference: Attitudes to yod in lexical and social contexts. *Journal of Greek Linguistics* 6: 151-185.
- Selkirk E. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (eds.) *Language sound structure: studies in phonology presented to Morris Halle by his teacher and students*, 107-136. Cambridge, MA: MIT Press.
- Setatos M. (1974). *Φωνολογία της Νέας Ελληνικής Κοινής*, Αθήνα: Παπαζήση.
- Smith J. L. (2005). *Phonological Augmentation in Prominent Positions*, New York and London: Routledge.
- Trudgill P. (2003). Modern Greek dialects: a preliminary classification. *Journal of Greek Linguistics* 4: 45-64.
- Trudgill P. (2009). Greek Dialect Vowel Systems, Vowel Dispersion Theory and Sociolinguistic Typology. *Journal of Greek Linguistics* 9: 165-182.
- Uffmann C. (2007). Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29: 451-76.
- Warburton I. (1976). On the boundaries of morphology and phonology: A case study from Modern Greek. *Journal of Linguistics* 12: 258-278.
- Yip M. (2003). Some real and not-so-real consequences of comparative markedness, *Theoretical linguistics* 29.1/2: 53-64.

Diminutive suffixes in modern greek dialects

GIANNOULA GIANNOULOPOULOU
University of Athens
giannoulop@isll.uoa.gr

1. Introduction

Suffixation by diminutives in Modern Greek and in its dialects belong to the morphological processes that occupy a central role in recent linguistic research, as far as both its morphological features (cf. Symeonidis 1968, Babinotis 1969, Koutita-Kaimaki 1984, Melissaropoulou & Ralli (in press), Ralli & Melissaropoulou 2007, Melissaropoulou 2009) and its pragmatic features are concerned (cf. Daltas 1985, Sifianou 1992).

The aim of the present paper is to present and analyse the diminutive suffixes in certain Modern Greek dialects and, more specifically, to examine whether the distinction between Northern and Southern dialects is accompanied by a relative differentiation as far as diminutives are concerned. The research continues a previous one (Giannoulopoulou 2006) about the different occurrence of compounding in Northern and Southern dialects, in which it was confirmed that compounds appear more frequently in Southern dialects than in Northern ones. That confirmation has been related with the syntheticity / analyticity features of the Modern Greek dialects.

Lexical units from ten (10) glossaries of Modern Greek dialects, which represent the distribution in Northern and Southern dialects are examined in the present study. More specifically, the Northern dialects of Agiasos (Lesvos), Veroia, Litochoro, Kozani and Pelion, the Southern dialects of Helia (Peloponnese), Zante, Xiromero, Crete and the Southern-east dialect of Pyrgi (Chios) are examined.

2. Frequency of diminutized lexical units in Northern and Southern Modern Greek dialects

The survey of suffixed by diminutive lexical units in the glossaries of dialects is a difficult process for two reasons: a) because not all diminutive formations are lemmatized in the glossaries, but just the ones that have a semantic salience; this happens not only with the glossaries but also with the linguistic vocabularies of Modern Greek and b) because the diminutive suffixes are lemmatized in just a few glossaries. In any case, glossaries are a useful source for the diminutive formation in dialects.

In a relatively extensive glossary of the Northern dialect of Veroia (2,552 words) just 9 diminutive formations are lemmatized, in which 5 different diminutive suffixes occur:

-aci, e.g. spaθ- 'sword', spaθ**aci** lit. 'little sword', metaphorical meaning 'lily', because the shape of the leaves resembles that of a sword,
-uði, e.g. litury- 'cake', litur**yuði** 'little cake which was given to the kids',
-itsa, e.g. kap- 'capote', kapi**nitsa** 'little capote',
-uli, e.g. kumats- 'piece', kumatsi**uli** 'little piece',
-iði, e.g. skaf- 'tub', ska**fiði** 'little tub for the preparation of bread'.

In an equal-sized glossary of the Southern dialect of Zante (1,716 words) 61 diminutive formations are lemmatized, in which 7 different diminutive suffixes are found:

-aci, e.g. anem- 'spinning-wheel', anemi**δaci** 'little spinning-wheel',
neut. -uði, fem. -uða, e.g. maθit- 'pupil', maθi**tuði** 'little boy who follows the priest',
vosk- 'stay in the same place', voska**ruða** 'bird that stays in the same place',
-itsa, e.g. pa**pitsa** 'iron',
neut. -uli, -fem. ula, e.g. yats- 'cat', ya**tsuli** 'little cat',
bal- 'ball', ba**lula** 'small ball',

-iði, e.g. xe- derivational prefix, cip- 'garden', xeci'piði 'remains of the harvest',
-opulo, e.g. cera'topulo, used to express admiration to kids,
-eli, e.g. kampan- 'bell', kampa'neli 'small bell'.

Comparison of the two glossaries shows that, although there is a difference between the two glossaries as far as the number of diminutive formations is concerned (9 / 61), this does not necessarily imply a similar difference in the number of the diminutive suffixes that are found in each case (5 / 7).

In the glossary of the Northern dialect of Agiasos (Lesvos) (2,700 words), where 62 diminutive formations are lemmatized, 4 different diminutive suffixes are found:

neut. -uði, fem. -uða, e.g. kupil 'girl', kupi'luð 'little girl',
neut. -uli, fem. -ula, e.g. krivats- 'bed', kriva'tsul 'small bed',
babak- 'cotton', babakula 'cotton thread',
-iði, e.g. akli'sið 'small church', akli- 'church'
-eli, e.g. sts'lupsar- 'shark', sts'lupsa'rel 'small shark'.

It is worth mentioning that in the glossaries of all the dialects there are found diminutive formations with the suffix -iði, which do not occur in the Standard Modern Greek.

E.g. akli- 'church', akli'sið 'small church' (Lesvos),
amps- 'nephew', am'psið 'little nephew' (Kozani),
skaf- 'tub', ska'fiði 'little tub for the preparation of bread' (Veroia),
xe- derivational prefix, cip- 'garden', xeci'piði 'remains of the harvest' (Zante),
kofin- 'basket', ksekofi'niði 'narrow and long basket' (Chios).

As shown from the above examples and as found in all the examined glossaries, it does not exist a remarkable difference between Northern and Southern dialects of Modern Greek as far as the productivity of diminution is concerned, regardless of the diminutive suffixes that are found in each case.

In a previous research (Giannouloupoulou 2006) on the differentiation between Northern and Southern dialects as far as compounding is concerned, it was shown that there does exist a prevalence of compounding in the Southern dialects compared to the Northern ones and this prevalence is connected with the growing analyticity of the Northern dialects compared to the Southern ones.

In the study of diminution a similar difference is not observed. This is not contradictory with the growing syntheticity of the Southern dialects and the growing analyticity of the Northern ones, because diminution is a process that takes place in derivation in Modern Greek (cf. Melissaropoulou & Ralli 2008, Karra 2006). This means that diminution concerns the co-existence of a lexical and a grammatical morpheme and not the co-existence of lexical morphemes and the concomitant syntheticity.

In certain cases of the examined data, diminutive suffixes function more grammatically than pure derivative suffixes, that is, they assign neither the purely diminutive nor the connotative affective meaning to the base of the word, but they function just as a marker of the class, they function as morphemes that enlist a word in the system just like the inflectional morphemes. E.g. The lexical unit *xasuli* (from the dialect of Veroia) meaning 'unripe cane of grain', which does not refer to something small, but the suffix -uli adjusts the Turkish loan word *hasil*.

Also worth mentioning is the lexical unit *γaδuli* (from the dialect of Zante), which has the unpredictable meaning 'big bucket', while there also exists the word *γaði* meaning 'bucket with holes'. This particular function of the diminutive suffixes is frequent in the adaptation of loan words and is also noticed in Standard Modern Greek. E.g. the lexical unit *bar* 'bar' is not adapted in the inflectional system of Modern Greek via another inflectional morpheme, but by means of the diminutive suffix -aci. The word *baraci* 'bar / little bar' is not different from the word *bar* as far as the size is concerned, but in the intimacy that the word assigns in the whole utterance.

Thus, it is observed that apart from their diminutive and affective meaning, diminutive suffixes also have another use by speakers as a strategy of adaptation in the system.

3. Do certain dialects have certain diminutive suffixes?

The prevalence of certain diminutive suffixes in certain dialects is an often-referred phenomenon in the literature. E.g. it is known that in Italian the suffix *-ino* prevails in the dialect of Toscana, while the suffix *-etto* prevails in the dialect of Venice.

In the Greek literature the suffixes *-uði* and *-eli* are considered as restricted in Macedonia and Lesvos (Dietrich 1928: 138-9), the suffix *-akos* restricted in Mani, while the common *-aci* is considered to have a special presence in Crete, where it is used for the formation of the family names in *-acis*. Correspondingly, the suffix *-opulos* is considered (cf. Dietrich op.cit.: 155) as a specific feature of family names in Peloponnese.

Prevalence of certain diminutive suffixes in certain dialects is confirmed in the data of the present research, but this does not mean that certain diminutive suffixes are excluded from certain dialects. E.g. the suffix *-eli* is found in the Southern dialect of Zante perhaps because of Italian influence: *kampaneli* 'small church bell', *katsuriðeli* 'small tree branch' diminutive of *katsuriða* 'long tree branch', *kurtunelia* 'bed curtains'. The same happens in Zante dialect with *-uði*: *apofauðia* 'food remains', *voskaruða* 'bird that stays in the same place'.

But both these suffixes are more frequent in Northern dialects. More specifically, the suffix *-eli* is extremely frequent in Lesvos.

The case of the suffix *-opulo*, which is more frequent in the Southern dialects, is a similar one. The original meaning of the suffix was patronymic and it was found in neuter gender, with the meaning 'offspring of humans or animals' (Dietrich, op.cit.: 154). Its expansion to bases meaning something inanimate rendered it a suffix with generalized diminutive meaning. This expansion is obvious especially in Southern dialects: *skia'dopulo* 'a sort of grape' (Zante), *kaðopula* 'small bucket', *masto'ropulo* 'young craftsman', *porto'pula* 'small door' (Akarnania).

But there are some exceptions. While in the data from Northern dialects suffixation with *-opulo* is rare, in the Northern dialect of Kozani several diminutivized with *-opulo* words are found: *ðimu'noplu* 'naughty boy', *ciara'toplu* 'naughty boy', *paraði'roplu* 'small window', *spi'toplu* 'small house', *klistoi'pula* 'young fighters'.

Certain suffixes are not exclusively connected with certain dialects and this is probably due to the fact that the glossaries of the present research are recent products and, thus, the influence of Standard Modern Greek to the dialects is strong. Probably, too, that has never been so.

It is also worth mentioning that in the dialect of Zante there exist loan suffixes from Italian. There are found lexical units as *γatsulinos* 'cat, small dogfish', *biskurðini* 'small delicacy' < Ital. dim. *-ino*, *bo'tsoni* 'small bottle' (*botsa* 'bottle'), *portoni* 'iron door', *stra'toni* 'small and narrow door' < Ital. augm. *-one*.

Such suffixed words are not found in the Northern dialects. There are also found in some Southern dialects, e.g. in Helia the words *stra'toni* and *por'toni* with the same meaning as in Zante, in Akarnania the word *stra'toni* meaning 'half an acre of a vineyard'.

Although most of these words are analyzable, namely their base is a theme of a free word in the dialect, these suffixes are not expanded in many themes of Greek words. In other words, the situation is in the boundaries between loaning of words and loaning of suffixes.

4. Accidental sequences of phonemes or reanalysis?

In the examined data several lexical units are noticed to end to sequences of phonemes which coincide with certain diminutive suffixes, without the possibility to separate the base from the suffix and to recognize morphological and semantic boundaries of suffixation. Most of these cases are loan words that end to *-itsa* and *-aci*. Linguistic research has been particularly occupied with the suffix *-itsa* from an etymological point of view and more specifically with its Slavic or Greek origin (cf. Georgakas 1982).

Anastassiadi-Symeonidi (1994: 205) states for the suffix -aci: “the affixoid bit -aci functions as a marker of incorporation, namely it is used in order to incorporate a non-adapted loan noun to the class of nouns in -aci. The element that plays a similar role in Corbin’s model, which has been applied in French, is called *intégrateur paradigmatique*”.

Our data present a complicated co-existence of diminutive suffixes with “accidental” sequences of phonemes, which has to be explained. E.g. in the dialect of Veroia there exist side by side suffixed words in -aci, (such as *spa’θaci* ‘little sword’ with the meaning ‘lily’, *pi’naci* ‘plate’) and words in -aci such as *tsiar’δaci* ‘small cottage’, *va’raci* ‘very slight piece of paper’. It is obvious that in this second category the element -aci functions as a marker of class, which incorporates in the Greek linguistic system the loans from Turkish in -ac. My proposal is that the co-existence with the diminutive suffix -aci urges speakers to morphological reanalysis of the adapted loans and to a gradual assignment of diminutive features to these.

By morphological reanalysis in the framework of Grammaticalization is meant “a new way in which speakers understand the structure of a word by relating it to other words in a different, novel way” (Haspelmath 1994: 1).

From the point of view of morphopragmatics in the above examples we may discern traces of meaning that are typical of diminutive suffixes. The morphopragmatic approach is proposed by Dressler & Barbaressi (1989, 1994) and consists in the incorporation of pragmatic meanings in the morphological rules. The study of diminutives in Italian has been fruitful for the development of morphopragmatics. Crocco-Galeas (2002: 153) shares the same point of view and assigns to the diminutive suffixes the following allo-pragmatic meanings: “1. Ludic character, 2. Meiosis, 3. Diminutivum puerile, 4. Child/lover/pet-centred speech situations, 5. Emotivity. 6. Familiarity and intimacy, 7. Sympathy and empathy”.

In our examples, the word *varaci* means ‘very slight piece of paper’, namely it is close to the diminutive meaning, while the word *tsiar’δaci* means ‘cottage’, namely it contains pejorative connotation.

Reanalysis has a pragmatic starting point. Speakers reanalyze by assigning diminutive connotative meaning to the sequence of phonemes -aci, since the majority of the words in -aci are diminutives. It is also possible that reanalysis obtains morphological status. In the case of *tsiar’δaci*, it is attested the word *tsar’δi*.

Similar observations can be made for the lexical units in -itsa, although the suffix -itsa has a complicated etymology. Some researchers –among them Chatzidakis and Andriotis–state that the suffix -itsa is a loan suffix from Slavic, where the suffix -ica is andronymic and in Greek is rendered a diminutive one. Others researchers, such as Koukoules and Georgakas, state that the suffix -itsa comes from the Medieval Greek suffix -icion. Regardless of the etymology, lexical units of Slavic origin with the suffix -itsa are found in the dialectal data. E. g. in the dialect of Kozani there is the word *gusta’ritsa* ‘big green lizard’ < Slav. *Guesteritsa*, side by side with the word *gustiaras*, which is coined with the supposed theme and the augmentative suffix -aras. In the same dialect the word *virvi’ritsa* ‘squirrel’ is found, which is probably of Slavic origin, and which in the dialect gains the metaphorical meaning ‘charming woman’ not only because of the referent ‘squirrel’ but also because of the diminutive suffix and its morphopragmatic connotations.

5. Conclusions

The following conclusions can be drawn from the above research:

- a) Diminution is an extremely productive derivative process in the dialects of Modern Greek, where both diminutive suffixes of Standard Modern Greek and special derivative dialectal suffixes are found.
- b) There is not to be observed a significant difference between the Northern and the Southern dialects as far as the productivity of diminution is concerned, in contrast with the observations that had been made in compounding.

- c) There do exist some diminutive suffixes which are typical of certain dialects, but these are not excluded from other dialects.
- d) In cases of co-existence of diminutive suffixes and homophone sequences of phonemes, especially in loan words, reanalysis of the loan words and assignment of diminutive features to the sequences are observed.

Further research of diminution in dialects will be useful for the study of the autonomy of this derivative process and for the development of morphopragmatics as a sub-discipline of morphology especially in languages, as Modern Greek, with rich derivative and inflectional morphology.

References

- Anastasiadi-Symeonidi A. (1994). *Νεολογικός Δανεισμός της Νεοελληνικής*. Thessaloniki.
- Babibiotis G. (1969). *Ο δια συνθέσεως υποκορισμός*. Athens: University of Athens.
- Crocco-Galeas G. (2002). A morphopragmatic approach to Greek diminutives, *Γλωσσολογικές έρευνες για την ελληνική*. L' Harmattan, Paris v. I., 151-154.
- Daltas P. (1985). Some patterns of variability in the use of diminutive and augmentative suffixes in spoken Modern Greek Koine (MGK). *Glossologia* 4: 63-88.
- Dietrich K. (1928). *Die Suffixe im Neugriechischen*. Balkan Archiv, 104-167.
- Dressler W. U. & Barbaressi L. (1989). Grammaticalizzazione morfopragmatica. In M. Berretta *et al.* (ed.). *Parallela* 4: 135-45. Narr, Tübingen.
- Dressler W. U. & Barbaressi L. (1994). *Morphopragmatics: Diminutives and Intensifiers in Italian, German and Other Languages*. Mouton de Gruyter, Berlin.
- Georgakas D. (1982). *A Graeco-Slavic controversial problem reexamined: the -ίτς- suffixes*. Athens: Publications of the Academy of Athens, 1107-6984.
- Giannouloupoulou G. (2006). Dialectological research and linguistic theory: the case of compounding. In *Proceedings of the 2nd International Conference of Modern Greek Dialects and Linguistic Theory*. Patras: University of Patras, 68-82.
- Haspelmath M. (1994). "The Growth of Affixes in Morphological Reanalysis". *Yearbook of Morphology*, 1-29.
- Karra, A. (2006). *Προβλήματα κλίσης και υποκορισμού στην ελληνική και στα ευρωπαϊκά συστήματα*. Unpublished Master's Thesis. Patras.
- Koutita-Kaimaki M. (1984). *Ο υποκορισμός στην ποντιακή διάλεκτο*. Thessaloniki: Aristotle University of Thessaloniki.
- Melissaropoulou D. (2009). Augmentation vs. Diminution in Greek Dialectal Variation: An Optimal System. In F. Montermini, G. Boyé, J. Tseng (eds.) *Selected Proceedings of the 6th Décembrettes*. Somerville, MA: Cascadilla Proceedings Project, 125-137.
- Melissaropoulou D. & Ralli A. (2008). Headedness in diminutive formation: Evidence from Modern Greek and its Dialectal Variation. *Acta Linguistica Hungarica* 55: 183-204.
- Melissaropoulou D. & Ralli A. (in press). Optimization in Grammar: A Contrastive Analysis of Diminutive Suffixation in Standard Modern Greek and the Asia Minor Dialect of Aivali. In *Proceedings of the 7th International Conference on Greek Linguistics* (York, 8-10 September 2005).
- Ralli A. & Melissaropoulou D. (2007). Ο Υποκορισμός στη Διαλεκτική Ποικιλία της Κοινής Νεοελληνικής. In *Proceedings of the 4th European Conference of the European Society of Modern Greek Studies*. Ellinika Grammata, Athens, 569-81.
- Reiner F. (1989). Appunti sui diminutivi italiani in -etto e -ino. In M. Berretta *et al.* (eds.) *Parallela* 4: 207-218. Narr, Tübingen.
- Sifianou M. (1992). The use of diminutives in expressing politeness: Modern Greek versus English. *Journal of Pragmatics* 17: 155-73.
- Symeonidis C. (1968). *Τα υποκοριστικά σε -κκο και -κκος της Καππαδοκικής*. Athens, Committee of Pontic Studies.

Glossaries

- Christodoulou C. (2003). *Τα Κουζιανιώτ'κα*. Thessaloniki.
- Davanos N.. *Λιτουχουρνή Ντουπιουλαλιά*. Litochoro: Community of Litochoro.
- Kanellakopoulos D. (2000). *Γλωσσάριο Πελοποννησιακής διαλέκτου, βασισμένο στην τοπική διάλεκτο του χωριού Λαδικού Ολυμπίας, νομού Ηλείας και των γύρω κοινοτήτων*. Athens.
- Konomos D. (1960). *Ζακυνθινό λεξικό*. Athens, 2003.

- Liapis K. (1996). *Το γλωσσικό ιδίωμα του Πηλίου*. Ores, Volos.
- Pagalos G. E. (1994-2002). *Περί του γλωσσικού ιδιώματος της Κρήτης*. Athens.
- Papanis D. & Papanis I. (2000). *Λεξικό της Αγιασώτικης Διαλέκτου*. Mytilene.
- Papatrechas G. (2007). *Ιδιωματικό και λαογραφικό γλωσσάριο Ξηρομέρου Ακαρνανίας*. Peristeri.
- Svarnopoulos, S. (1973). *Γλωσσάριο της Βέροιας*. Veria.
- Tsikis, N. (2002). *Γλωσσικά από το Πυργί της Χίου*. Athens.

Cypriot anomalies in *wh*-in situ structures*

KLEANTHES K. GROHMANN
University of Cyprus
kleanthi@ucy.ac.cy

ELENA PAPADOPOULOU
University of Essex
epapadb@essex.ac.uk

1. Introduction

This paper investigates structures of *wh*-in situ in (and corresponding interpretations available to speakers of) Cypriot Greek, a typical *wh*-ex situ language. That is, in order to form *wh*-questions, a single *wh*-phrase is fronted into the left sentence periphery, as in English, but under certain (pragmatic, discourse-specific) conditions, as in English, in-situ *wh*-expressions are felicitous to form an information question (i.e. without echo or rhetorical interpretation).¹ What makes Cypriot Greek potentially interesting in this respect is that, from all we know about its grammar — admittedly, not as much as we would like to —, structures that should not be possible or should be less preferred than others seem to be used and interpreted (and vice versa), in particular when compared to the closely related standard variety of Modern Greek. We set out to investigate some such structures quantitatively by conducting a questionnaire-based study on both syntactic structures and available interpretations of *wh*-in situ in Cypriot Greek.

To provide a very basic background to the language(s) discussed here, Cypriot Greek (henceforth, CG) is a linguistically understudied variety of Standard Modern Greek (henceforth, SMG) spoken on the island of Cyprus, in the far east of the Mediterranean Sea (more than twice as far from Athens as Rhodes, one of the southeastern-most islands of Greece). Several politico-economic reasons as well as an “apparent inability” of native speakers to draw linguistic boundaries between CG and SMG have led to a confusion as to what is “purely dialectal” and what “grammatically correct” means.² Opposing views regarding how similar or different the syntax of the two varieties is (Papagelou 2001) have guided a growing body of research carried out in Cyprus and elsewhere (e.g., Grohmann *et al.* 2006 and Gryllia & Lekakou 2006 on *wh*-related issues). More fundamental issues are currently being investigated for child language development by the Cyprus Acquisition Team (Grohmann, to appear, and much ongoing work).

We tested CG-speaking adults for interpretive effects in matrix and embedded information-question environments for *wh*-in situ vs. *wh*-ex situ and found a number of remarkable properties. One obvious factor in the licensing of such questions is the

* This paper started out as a reaction to Christos Vlachos’ presentation at the UCY Linguistics Discussion Group on *wh*-in situ in Greek (subsequently written up as Vlachos 2008, but revised as Vlachos 2010, with differences we will address here in some detail). We would like to thank Christos for discussion as well as Marcel den Dikken, Terje Lohndal, Panos Pappas, and the other reading group members, also for initial (dis)confirmation of judgements, in particular: Anna Epistithiou, Skevi Hadjiefthymiou, Evelina Leivada, Skevi Mavroudi, Chrystalla Michael, Natalia Pavlou, and Elena Theodorou. We extend our gratitude to the audiences at the *ISTAL 19* workshop ‘The Optionality of Wh-Movement’ (Thessaloniki, April 2009) and the *MGDLT 4* conference (Chios, June 2009). A revised version of this paper is going to appear as Grohmann & Papadopoulou (forthcoming).

¹ For reasons of simplicity, the discussion is restricted to single information questions throughout this investigation. The major points to be highlighted hold irrespective of the number of *wh*-expressions. On the theoretical relevance of echo questions for minimalist analysis, see the very recent work by Sobin (2010), whose relevance for the present topic we discuss in Grohmann & Papadopoulou (forthcoming).

² The use of double quotes here is intended to signal the difficulty researchers are faced when investigating a mostly “dialectal sub-standard” variety such as CG which is often flat out rejected by its own speakers as a “proper language” (see also Papapavlou 1998 and much subsequent work on socio-linguistic aspects of CG, and a current survey by Ioannidou *et al.* 2009, but see fn. 18 below).

contextual information in a way yet to be described appropriately (see also the references in the following paragraph), but in the long run we aim to incorporate a better developed effect of the role and use of (CG) discourse-linked questions as opposed to single *wh*-expressions as done in this study.

Our contribution explores how the variety of Greek spoken on the island of Cyprus differs in interesting ways from mainland Greece. *Wh-in situ* in Greek, a *wh*-movement language, is discussed in section 2 (see Sinopoulou 2009 for SMG) alongside *wh-in situ* in English, another typical *wh*-movement language (cf. Ginzburg & Sag 2000). A very basic description of the phenomenon is presented in section 3, returned to in section 5 (based largely on Vlachos 2008, 2010). The discourse contexts in which *wh-in situ* is felicitous are presumably identical for SMG and CG, even possibly English and beyond (though they are not discussed here) — but the syntactic operations involved and semantic interpretations available are (or at least, may be) not. This is discussed at length empirically in section 4. Section 5 is the theoretical core of the paper that concludes the study with an extended analysis, discussion, and outlook.

2. *Wh*-Question Formation in CG

This study investigates the relationship between four types of *wh*-questions in Greek, those involving *wh*-arguments, such as *pjos/pcos* (SMG/CG) ‘who-MASC.NOM’ for subject and *pjon/pcon* (SMG/CG) ‘who-MASC.ACC’ for object as well as the manner-adjuncts *pos* and *indalos*, both meaning ‘how’ (in this section and the next). CG *wh*-question formation resembles to a large extent *wh*-question formation in SMG but differs with respect to some properties carried by CG *wh*-words and the addition of the dialectal element *embu* (Grohmann *et al.* 2006), literally ‘(it-)is-that’ (CG, as SMG, is a null-subject language); since it is used here in interrogatives (for non-interrogative focus use, see Fotiou 2009), we consider *embu* as ‘is(-it)-that’. However, as can be inferred from the results of the questionnaire complementing the study (section 4), more substantial differences arise (sections 3 and 5).

To set the stage for the structures to be discussed presently, (1) and (2) illustrate (regular) *wh-ex situ* and (specially conditioned) *wh-in situ* information questions with *wh*-arguments in SMG and CG.

- | | | | |
|-----|----|---|-------|
| (1) | a. | Pja/Pjo koritsi sinantise o Nikos xθes vraði? | [SMG] |
| | | who/which girl met the Nick yesterday evening
‘Who/Which girl did Nick meet last night?’ | |
| | b. | Pcan/Pcan koruan ivren o Nikos extes ti nixta? | [CG] |
| | | who/which girl found the Nick yesterday the night
‘Who/which girl did Nick meet last night?’ | |
| (2) | a. | O Nikos sinantise pja/pjo koritsi xθes vraði? | [SMG] |
| | | the Nick met who/which girl yesterday evening
‘Nick met who/which girl last night?’ | |
| | b. | O Nikos ivren pcan/pcan koruan extes ti nixta? | [CG] |
| | | the Nick found who/which girl yesterday the night
‘Nick met who/which girl last night?’ | |

CG *wh*-words bear an obvious morphological resemblance to their SMG counterparts, other than the obvious (and minor) morpho-phonological differences. The *wh*-expressions include the quantifiers *pcos/-ia/-o* ‘who/which’, *posos* ‘how much/many’, *ti* ‘what’, and *inda* ‘what’ as well as the adverbs *pote* ‘when’, *pu* ‘where’, *jati* ‘why’, *pos* ‘how’, *inda* ‘why’, and *indalo(i)s* ‘how’ (Simeonidis 2006:217; cf. Holton *et al.* 1997:414 for SMG). The quantifier *inda* ‘what’, and the adverbs *inda* ‘why’ and *indalos* ‘how’, are dialect-specific to CG as depicted in Table 1, which lists simplex *wh*-expressions in the left and (corresponding) complex ones in the right column.

Table 1: *Wh-words in Cypriot Greek*

Wh-quantifiers	
<i>pc-os/-ia/-o</i> 'who-MASC/-FEM/-NEUT'	
<i>pc-os/-ia/-o</i> NP 'which-MASC/-FEM/-NEUT NP'	<i>se pcon</i> 'to whom' <i>apo pcon</i> 'from whom' <i>*pu pcon</i> 'from whom' <i>pros pcon</i> 'to whom'
<i>pos-os/-in/-o</i> 'how much-MASC/-FEM/-NEUT'	<i>ja poso</i> 'for how long' <i>se poso</i> 'in how long' <i>*pu poso</i> 'from how much'
<i>ti</i> 'what'	<i>se ti</i> 'to what' <i>apo ti</i> 'from what' <i>pros ti</i> 'why'
Wh-adverbs	
<i>pote</i> 'when'	<i>apo pote</i> 'since when' <i>mexri pote</i> 'until when' <i>ja pote</i> 'for when'
<i>pu</i> 'where'	<i>apo pu</i> 'from where' <i>pros ta pu</i> 'towards where' <i>ja pu</i> 'to where'
<i>jati</i> 'why'	
<i>pos</i> 'how'	
CG-specific	
<i>indalo(i)s</i> 'how'	
<i>inda</i> 'what'	<i>se inda</i> 'in which' <i>pu inda</i> 'from which' <i>gia inda</i> 'for what'
<i>inda</i> 'why'	

*These are also specific to the CG dialect.

According to Simeonidis (2006:217), the CG *wh*-quantifier *inda* derives from the interrogative pronoun *tinda* 'what' used in Asizes (a text of laws from the island dating to the 10th and 11th centuries), literally *ti ine afta* 'what are these'.³ CG *inda* is a pronoun invariant in gender, number, and case which can be used either pronominally ('what/which NP') or pronominally (what we also call "bare *inda*" meaning simply 'what'). In addition, *inda* has the two phonologically reduced forms *a* and *nda*, which are used rarely and mainly in the village variety of the dialect known as "*xorkatika*" (Newton 1972:19). However, *inda* can also mean 'why' in CG, suggesting that this *inda* must have originated from *gia inda logo* 'for what reason' (Papadopoulou, in progress). When adjoined to (*e*)*mbu* 'is(-it)-that', both instances of *inda* come in several variants, namely, *nambu*, *tambu*, *ambu*, and *innambu* (Pavlou 2009, this volume). The third *inda*-derived *wh*-word is *indalo(i)s* 'how', literally *inda* 'what' + *logis* (in Ancient Greek *tropos*) 'way/manner', meaning 'in what manner, how', which also originated from the interrogative pronoun *tinda* 'what' (see e.g. Papagelou 2001, Simeonidis 2006, and Giagoulis 2009 for more discussion).

These three *inda-wh*-words have different properties from their SMG counterparts. SMG *pos* 'how', as in (3), can undergo movement into the left periphery (to a landing site

³ As Angeliki Ralli mentions in her state-of-the-art review article on Greek dialects, Contossopoulos (1983-1984), "who tries to establish an isogloss on the basis of the form of the *wh*-word *what*" (Ralli 2006:138), could also be cited here for work on dialectal question formation in Greek and the issue of *inda* (vs. *ti*).

one may assume to be Spec-C) or it can be left in situ (possibly adjoined to *v*/VP; see also section 3). When in situ, *pos* carries a more “restrictive” reading in SMG (as Vlachos 2008 calls it); the dialectal counterpart *indalos* ‘how’ does not share that property, since it can only appear sentence-initially (Papadopoulou, in progress), shown in (4). (We will return to these readings in section 3, and then again, more analytically, in section 5.)

- (3)a. **Pos** anikse tin porta o Nikos? [SMG]
 how opened the door the Nick
 ‘How did Nick open the door?’
 b. O Nikos anikse tin porta **pos**?
 the Nick opened the door how
 ‘Nick opened the door how?’
- (4)a. **Indalos** aniksen tin portan o Nikos? [CG]
 how opened the door the Nick
 ‘How did Nick open the door?’
 b. * O Nikos aniksen tin portan **indalos**?
 the Nick opened the door how
 ‘Nick opened the door how?’

Similar properties are exhibited by dialectal *inda* and SMG *jati* ‘why’, as well as CG *inda* and SMG *ti* ‘what’. On the ‘why’ side, SMG *jati* can either undergo movement to Spec-C or remain in situ, as in (5) below, whereas *inda* can only undergo movement, as in (6). Only if preceded by *ja* ‘for’ can *inda* be left in situ, as (6c) shows; in this environment, *inda* is freely translated as ‘why’ but literally should be, as glossed, ‘for what (reason)’.⁴

- (5)a. **Jati** piye ston ayona o Nikos? [SMG]
 why went to-the match the Nick
 ‘Why did Nick go to the match?’
 b. O Nikos piye ston ayona **jati**?
 the Nick went to-the match why
 ‘Nick went to the match why?’
- (6)a. **Inda** epien is tin mappan o Nikos? [CG]
 why went to the match the Nick
 ‘Why did Nick go to the match?’
 b. * O Nikos epien is tin mappan **inda**?
 the Nick went to the match why
 ‘Nick went to the match why?’
 c. O Nikos epien is tin mappan **ja inda**?
 the Nick went to the match for what
 ‘Nick went to the match why?’

Prenominal *inda* ‘what’ can remain in situ or undergo movement in both SMG and CG, shown in (7) and (8), respectively.

- (7)a. **Ti vivlio** ðiavazi o Nikos? [SMG]
 what book reads the Nick
 ‘What book is Nick reading?’

⁴ A more detailed analysis of these structures is provided by Pavlou (this volume) and Papadopoulou (in progress). Our main concern here regards (non-)availability of *wh*-in situ in CG and the corresponding interpretations as well as purported “mismatches” or unexpected structures, discussed from section 3 on.

- b. O Nikos *ðiavazi ti vivlio?*
the Nick reads what book
'Nick is reading what book?'

(8)a. **Inda vivlion** *θkiavazi o Nikos?* [CG]
what book reads the Nick

'What book is Nick reading?'

- b. O Nikos *θkiavazi inda vivlion?*
the Nick reads what book
'Nick is reading what book?'

Pronominal or bare *inda* 'what', on the other hand, obligatorily undergoes movement to Spec-C and can never be left in situ, as (10) demonstrates (more on *(e)mbu* below), in contrast to *ti* (predominantly used in SMG, but also employed by CG speakers), shown in (9).

(9)a. **Ti** *ðiavazi o Nikos?* [SMG & CG]
what reads the Nick
'What is Nick reading?'

- b. O Nikos *ðiavazi ti?*
the Nick reads what
'Nick is reading what?'

(10)a. **Indambu** *θkiavazi o Nikos?* [CG]
what-EMBU reads the Nick
'What is Nick reading?'

- b. * O Nikos *θkiavazi indambu?*
the Nick reads what-EMBU
'Nick is reading what?'

Notice that bare *inda*, i.e. when used pronominally, is always followed by *mbu*, which arguably is a phonological variant of *embu* 'is(-it)-that' (Grohmann *et al.* 2006). Promising accounts would take bare *inda* to have grammaticalized as *indambu* 'what-is(-it)-that' (Papadopoulou, in progress) or perhaps combine with it syntactically (see Pavlou, this volume, for discussion of several possibilities); we assume the former.⁵ For readability, we often gloss *(e)mbu* 'EMBU'.

A characteristic property of CG *wh*-question formation is the addition of this element *embu* which may optionally appear after the preposed *wh*-word, deriving questions such as (11a) and (12a) below. Depending on how *embu* is analyzed, different syntactic operations would be involved in the derivation of CG *wh*-questions. Initially (cf. fn. 5), it was suggested that *embu*-structures are essentially *bona fide* cleft-structures (Grohmann *et al.* 2006), but considering that SMG does not allow any form of clefting, such a syntactic innovation may be a little far-fetched, so that *embu*-structures might rather involve a "fossilized" complementizer, where interrogative C be filled by *embu* (Papadopoulou, in progress).

Regardless of the final analysis of *(e)mbu*, the following data illustrate the (im)possibilities of *pos/indalos* 'how' in CG:

⁵ We leave aside the original suggestion by Grohmann *et al.* (2006), briefly alluded to in the text presently, that *embu* actually contains or introduces a full-fledged clefting structure, akin to English "It is X that..." (see Fotiou 2009 for non-interrogative focus but also Gryllia & Lekakou 2006 for some criticism), or the possibility they suggest but then reject that, when reduced to *mbu*, the *wh*-word *inda* undergoes *wh*-cliticization parallel to what may be found in Romance varieties (cf. Munaro & Pollock 2005).

- (11)a. **Pos** (*embu*) aniksen tin kashian o Nikos? [CG]
 how EMBU opened the box the Nick
 ‘How did Nick open the box?’
 b. O Nikos (**embu*) aniksen tin kashian **pos** (**embu*)?
 the Nick EMBU opened the box how EMBU
 ‘Nick opened the box how?’
- (12)a. **Indalos** (*embu*) aniksen tin kashian o Nikos? [CG]
 how EMBU opened the box the Nick
 ‘How did Nick open the box?’
 b. * O Nikos (*embu*) aniksen tin kashian **indalos** (*embu*)?
 the Nick EMBU opened the box how EMBU
 ‘Nick opened the box how?’

Note that *embu* ‘is(-it)-that’ cannot be found along with the *wh*-word in situ, even though the *wh*-word on its own can, as in (3b) and the b-examples of (5)–(9). (11b), in particular, shows two things: (i) CG-used *pos* may stay in situ, unlike CG *indalos* (cf. (12b)), and (ii) *embu* can neither occur in a low position near an in-situ *wh*-expression nor appear in the left periphery on its own. The ban on occurrences of *embu* in the clause can in fact be schematized as in (13).

- (13) a. [_{CP} WH ((*e*)*mbu*) ... t_{WH} ...]
 b. * [_{CP} ((*e*)*mbu*) XP ((*e*)*mbu*) ... WH ((*e*)*mbu*) ... ((*e*)*mbu*)]

More can and possibly should be said, but since the remainder of this paper will not deal with *embu* as such (see e.g. Grohmann *et al.* 2006, Fotiou 2009, and Papadopoulou, in progress), this characterization that *embu* is restricted to a left-peripheral position right-adjacent to a fronted *wh*-expression, bare and rough as it is, hopefully suffices. In other words, *embu* (or, as discussed in Grohmann *et al.* 2006 and, at length, Pavlou this volume, *mbu* when following variants of *inda* ‘what’ and ‘why’) is restricted to optional occurrence in an interrogative C.⁶

Other than the *embu*-strategy, the first major difference between SMG and CG *wh*-question formation, then, is that the native item for ‘how’, *indalos*, cannot stay in situ (as in (4b) and (12b)), unlike SMG, where *pos* may stay in situ (as in (3a)). The Greek form *pos*, when used by speakers in CG, is also allowed in situ (as in (11b)). The same holds for CG *inda* ‘why’ (cf. (6b)) and *inda(mbu)* ‘what(-EMBU)’ (cf. (10b)) as opposed to the corresponding SMG *jati* and *ti*, respectively, even when used in CG (cf. (5b) and (9b)).

To address *wh*-in situ non-reprising, information questions very briefly (beyond Bolinger 1978 and Ginzburg & Sag 2000), it is clear that they require a particular discourse context. Vlachos (2008, 2010) goes into significant detail in his general account of such structures in SMG and we do not think that much more needs to be said for the purposes of the present paper. We thus restrict ourselves to pointing out that intuitively, one of the facilitating factors involved seems to be something very much akin to D(iscourse)-linking (Pesetsky 1987), that is, in order to ask a *wh*-in situ question felicitously, a discourse context must have been established that allows identification of the *wh*-expression. Other than difficulties examples such as (15b) might bring about, this cannot be the whole story, however, as Vlachos (2010) also demonstrates, but it helps assigning an initial analysis of *wh*-in situ in terms of “unselective binding” (Kamp 1981, Heim 1982; see also e.g. Cresti 1998), as also suggested by Pesetsky for D-linking, under

⁶ Note that Grohmann *et al.*’s (2006) clefting-approach to *embu* can capture the distributional facts as well, since there *embu* is decomposed into copular *en* plus complementizer *pu* that “fuse” (post-)syntactically.

which the *wh*-expression would be bound by an interrogative operator; Vlachos proposes an alternative that licenses the in-situ syntax more locally, within the *vP*, and all we care about here, regardless of the specifics, is that in-situ *wh*-items can indeed be licensed in situ (see also the beginning of the next section).

Some examples of *bona fide* information questions with *wh*-in situ in English follow (Ginzburg & Sag 2000:280), some construed, others taken from the “real world” (English in-situ *wh*-expressions require special stress, indicated by small capitals):

- (14) a. A: Well, anyway, I’m leaving.
 B: OK, so you’ll be leaving when exactly?
 b. A: I’m annoyed.
 B: Aha. You’re annoyed with whom?
- (15) a. A: My friends, they saw everything.
 B: Yeah, they saw what?
 [CBS Saturday Night Movie, 25 January 1992]
 b. Michael Krasny [addressing a guest — who has not said anything yet — about the interim chief of the US Attorney’s office]:
 This is a position that is how important in your judgment, Rory?
 [Forum KQED, 29 July 1998]

Pending further discussion, an in-situ *wh*-item WH can be bound unselectively by a question operator OP (CP) or licensed locally (*vP*):

- (16) [_{CP} (OP_i) C_Q ... [_{vP} (OP_i) ... WH_i ...]]

3. Ex-Situ and In-Situ Interpretive Quirks

Aside from the variation in SMG and CG question formation so far discussed, stronger divergences arise regarding different restrictions in interpretation, that is, the kinds of readings speakers associate with in-situ structures. *Wh*-words left in situ do so at the cost of interpretation.

Generally, a *wh*-item is interpreted in its scope position or rather, it scopes over material *c*-commanded from its interpretation site. In ex-situ constructions, the *wh*-item thus scopes over the entire clause from its Spec-C position. A question that then arises for in-situ *wh*-constructions is what scope they take. Typical *wh*-in-situ languages such as Chinese are not restricted as such by clause boundaries, that is, an embedded in-situ *wh*-expression can take matrix scope (Huang 1982 and much subsequent work). Vlachos (2008, 2010) has shown for SMG that *wh*-in situ expressions are clause-bound.⁷ This section will address some pertinent issues for CG *wh*-in situ — and some possibly quite puzzling, astounding differences from SMG.

Before we go there, however, three remarks are in order. First, one may ask to what extent SMG *pos* (as well as *jati* and *ti*) used in CG would indeed reflect CG — or in other words: Can Greek words be used at all in the Cypriot dialect? Put this way, the answer must be a resounding “Yes”: After all, not every word of the CG variety is uniquely native. But the trickier part of this question is whether in this case two synonymous words can be said to be “in competition” — or whether they are either not synonymous after all or do not really compete. If they were not synonymous, we would not face an issue here, but

⁷ The relevant comparison would, of course, not be with a strict *wh*-in-situ languages, but with one that allows optional *wh*-in situ, such as French (argued to be clause-bound and restricted to root clauses starting with Chang 1997 and Bošković 1997). Much has been debated over the correct properties of French *wh*-in situ, and the upshot seems to be that there are at least two varieties (Mathieu 2004), one that allows and one that disallows embedded *wh*-in situ (see e.g. Starke 2001, Cheng & Rooryck 2002 and Adli 2006).

from the limited data we have gathered, we cannot discern whether this is indeed the case. Lack of competition could mean two things: The SMG form comes with SMG syntax, even when used in a CG context, or something else is going on.

Pending further discussion and digression, we assume that the use of SMG items in CG speech is not only acceptable, but also does not take away anything from the CG-specific grammatical properties under investigation. We also leave aside the issue whether idiolects, diglossia, and other sociolinguistic influences a “high” variety may have on a “low” one and follow standard generative assumptions that the language of a speaker is the result of an internalized grammar of that speaker — yes, “dialects” have their own grammar, on a par with “languages” (cf. Kayne 2000) — and if a large group of CG speakers employs *pos*, it reflects the clear availability of *pos* in that group’s lexicon rather than code-switching or any other “explanation” one might want to bring up. Variations of our answer to the first remark may also become clearer when we look at the third point raised below.

Second, it might be debatable at first sight whether the “in-situ” *wh*-items (in either variety) are indeed *in situ*. We will not engage in a discussion as to what the (arguably, predominantly discourse-driven) factors are that allow in-situ information questions, that is, the “non-reprising” use of in-situ questions, first observed by Bolinger (1978), more recently discussed by Ginzburg & Sag (2000: chap. 7). Vlachos (2008, 2010) does this at length in a modern, minimalist framework taking into account formal semantic and pragmatic notions. Rather, the question is meant to tie in “apparently in-situ expressions” with an analysis that assumes lower projections as landing sites for short (*wh*-) movement, as suggested recently by Belletti (2004), for example. The idea here is that discourse-related positions, such as topic and focus (and, by extension, *wh*-items), are not uniquely licensed in the clausal left periphery (“split Comp” in the sense of Rizzi 1997), but that they can also appear in the “lower Infl” area, such as at the periphery of *vP* or, to use current terminology, at the outer edge of the “*vP*-phase” (in Phase Theory, starting with Chomsky 2000). Sinopoulou (2008) applies this idea to Greek multiple *wh*-questions, but explicitly not to single *wh*-in situ (see also Sinopoulou 2009 and Vlachos 2008, 2010).

Again, we side with Vlachos (2010), who provides an interesting account in the context of the larger issues of *wh*-in situ, non-reprising information questions (see also the brief discussion around (16)) which might, in the end, be compatible with either view, depending on one’s take on displacement in natural language, but it does make a strong case for “in-situ *in situ*” as we assume here for simplicity. In addition, we hold the perhaps conservative view that different parts of the clause structure are responsible for different interface tasks — but uniquely so. Referring to the tripartite, domain-driven framework of Grohmann (2003), the lowest part of the structure is responsible for thematic information (an articulate *vP*, which he calls “ Θ -Domain”), while discourse-related material and operators must be licensed in the highest part (“split Comp” or an articulated CP, the “ Ω -Domain”), couching the agreement-layer in between (“split Infl” or an articulated TP, the “ Φ -Domain”). In other words, we assume a three-way split of clausal structure into CP – TP – *vP*, each expanded into different functional projections, but also each uniquely identifying interpretive tasks. This view does not easily allow low discourse-related licensing, unlike Belletti’s (2004) approach which, in turn, might be easily made compatible not only with the cartographic framework assumed there, but also, as briefly mentioned above, with Phase Theory in a perhaps natural manner.⁸ The long and short of the second remark, then, is that we assume the in-situ *wh*-phrase not to have moved at all, even if it may be only for convenience at this point.⁹

⁸ For a preliminary discussion on how to frame some of Grohmann’s (2003) core insights within Phase Theory, see Grohmann (in press).

⁹ A potential consequence might be that an unselectively binding operator from a CP-related position should be clearly preferred over a local *vP*-operator (cf. the very basic (16) above); we will not pursue this issue any further.

Third, and related to the previous point, the fact alone that two lexical items show different syntactic behavior is not that surprising — after all, they are different lexical items. Even in English, it has been argued that not all *wh*-items pattern alike. The “true adjuncts” *why* and *how*, for example, have been suggested to be generated high, inserted directly into C, unlike “(semi-) argumental” *who*, *what*, *when*, etc. (Rizzi 1990; see also Bromberger 1987 on English *why*, Collins 1991 on English *how come*, Ko 2005 and Ochi 2004 for valuable cross-linguistic discussion, and Tsai 2008 for more recent discussion). More relevant is the observation that the two *wh*-expressions for reason (but see Tsai 2008 for a more fine-grained distinction between *wh*-adjuncts which goes beyond the scope of the present investigation of CG), *why* and *how come*, show quite different properties within the same language — for example, *how come* does not trigger inversion and it may not stay in situ. In this respect, *pos* and *indalos* in CG might reflect *why* and *how come* in English, respectively.

With all this in mind, we suggest here that *indalos* is obligatorily merged into Spec-C (presumably specified as such in the CG lexicon), while *pos* at least may come from a lower position (leaving open the option of “high insertion” if it turns out to be needed). When doing so, scope ambiguities might arise — and should be resolved with in-situ *wh*-expressions. Consider the schematic structures in (17):

- (17) a. [CP *indalos* (*embu*) [...]]
 b. [CP *pos* (*embu*) [... *t_{pos}* ...]]

The high-inserted *indalos* obligatorily takes scope over the entire clause, while *pos* may at least in theory take the same “high scope” — but in addition also “low scope” if interpreted in its base position. The following data illustrate what we have in mind.

Take a simple English sentence like (18):

- (18) John opened the door.

At least two relevant modifications can be expressed, an *instrumental modification* (expressing the instrument with which the door was opened) or a *manner interpretation* (referring to the manner, or in this case better: disposition, of the agent of the door-opening event):

- (19) a. John opened the door with the key. *instrumental*
 b. John opened the door with anger/angrily. *manner*

A *how*-question gives rise to ambiguity: *How did John open the door?* could be answered with either (19a) or (19b). The same holds for Greek. In particular, as Vlachos (2008) first discussed, when the *wh*-expression is in Spec-C, both readings are available, as in (20).

- (20) **Pos** anikse tin porta o Nikos? [SMG]
 how opened the door the Nick
 ‘How did Nick open the door?’
 a. Me to kliði.
 with the key
 ‘With the key.’
 b. Nevriasmenos.
 angry-NOM
 ‘With anger.’

In contrast, in-situ *pos* only allows the instrumental interpretation:

- (21) O Nikos anikse tin porta **pos**? [SMG]
 the Nick opened the door how
 ‘Nick opened the door how?’
 a. Me to kliði.
 with the key
 ‘With the key.’
 b. # Nevriasmenos.
 angry-NOM
 ‘With anger.’

We will return in section 5, where we address additional factors and complications, to the at first glance puzzling fact that CG seems to differ in this respect along the lines of (24) below. CG *pos*, namely, seems to allow both interpretations in both situations, irrespective, thus, of whether the *wh*-word is in situ or not. That is, (21) is perfectly acceptable with a manner interpretation in CG, as (22) shows.

- (22) O Nikos aniksen tin porta **pos**? [CG]
 the Nick opened the door how
 ‘Nick opened the door how?’
 a. Me to kliði.
 with the key
 ‘With the key.’
 b. Nevriasmenos.
 angry-NOM
 ‘With anger.’

Similarly to CG *pos*, argumental ‘who’-questions allow different readings in more complex contexts (data again taken from Vlachos 2008). In SMG questions such as (23), both readings are available, where the *wh*-phrase can either be construed with the matrix (object of *anakinose*) or the embedded clause (as the argument of *apokalipse*).

- (23) **Se pjon** anakinose o Janis oti i Maria apokalipse to mistiko? [SMG]
 to whom announced the John that the Mary revealed the secret
 ‘To whom did John announce that Mary revealed the secret?’
 a. To anakinose ston diefθindi tu.
 it announced to-the senior-manager his
 ‘He announced it to his senior manager.’
 b. Anakinose oti i Maria to apokalipse ston adayonisti tis eterias.
 announced that the Maria it revealed to-the competitor of-the company
 ‘He announced that Mary revealed it to the competitor of the company.’

The two interpretations are arguably derived from a simplified structure, such as the one depicted in (24).¹⁰ That is, movement of the *wh*-word to Spec-C either from the

¹⁰ We only mark VP very broadly, not committing to the internal structure of ditransitive predicates and following standard assumptions that the verb moves at least to T. We also assume that the post-verbal subject preceding the predicate’s internal arguments stays in situ (Spec-v), whereas the pre-verbal subject position may either be Spec-T or some higher position, such as a topic phrase. The exact details, an issue of perennial debate in Greek syntax, do not play a role here; for discussion, see, among many others, Philippaki-Warbuton (1985), Alexiadou & Anagnostopoulou (1998), and Roussou & Tsimpli (2006).

This said, as pointed out to us by Spyros Armostis (p.c.), there is a mismatch which we unfortunately failed to control for in the quantitative data discussed in section 4: Note that the subject in the matrix clause is post-verbal, while in the embedded clause it shows up in pre-verbal position. To the extent that this might be relevant, we have not been able to integrate it into the

matrix (position $_A$) or from the embedded clause (position $_B$) allows it to be construed as the internal argument of the matrix or the embedded verb, respectively.

- (24) Se pjon anakinose o Janis [_{VP} $t_{\text{anakinose } _A}$
[oti i Maria apokalipse [_{VP} to mistiko $t_{\text{apokalipse } _B}$]]]?
'To whom did John announce that Mary revealed the secret?'

The same should apply in CG embedded *wh*-questions — but as signaled in (25), the embedded reading is marginal, if possible at all (see section 4 for quantitative results and section 5 for discussion, including the reason why we translate *esinaferen* as 'said').

- (25) **Se pcon (embu)** esinaferen o Yiannis oti i Maria ipen tin alithkian? [CG]
to whom EMBU talked-about the John that the Mary said the truth
'To whom did John say that Mary said the truth?'
- a. Ipen to ston Giogo.
said it to-the George
'He said it to George.'
 - b. # Ipen oti i Maria ipen stin Anna tin aliθkian.
said that the Maria said to-the Anna the truth
'He said that Mary said the truth to Anna.'

Restrictions similar to (21) above account for embedded in-situ *wh*-phrases in SMG. Assuming in-situ *wh*-phrases to be clause-bound in SMG (Vlachos 2008), they should not be able to be interpreted as an argument of the matrix verb. And indeed, in (26), the *wh*-phrase is interpreted as the argument of the embedded clause only, and not the matrix, allowing for the b- but not the a-interpretation (Vlachos 2010).

- (26) O Janis anakinose oti i Maria apokalipse to mistiko **se pjon**? [SMG]
the John announced that the Mary revealed the secret to whom
'John announced that Mary revealed the secret to whom?'
- a. # To anakinose ston diefθindi tu.
it announced to-the senior-manager his
'He announced it to his senior manager.'
 - b. Anakinose oti i Maria to apokalipse ston adayonisti tis eterias.
announced that the Mary it revealed to-the competitor of-the company
'He announced that Mary revealed it to the competitor of the company.'

Not so in CG, however. With a final *wh*-phrase, which we may take to be in situ within the embedded clause for now (but see section 5 for discussion), interpretation of the *wh*-phrase as the argument of either the matrix or the embedded clause is allowed, as in (27); again, see section 4 for speakers' judgments results from the questionnaire.

- (27) O Yiannis ipen oti i Maria esinaferen tin aliθkian **se pcon**? [CG]
the John said that the Mary talked-about the truth to whom
'John said that Mary said the truth to whom?'
- a. Ipen to ston Giogo.
said it to-the George
'He said that to George.'
 - b. Ipen oti i Maria ipen stin Anna tin aliθkian.
said that the Mary said to-the Anna the truth
'He said that Mary said the truth to Anna.'

discussion of this paper and leave it as an intriguing research question for the future.

It is apparent that CG and SMG do not differ only with respect to some dialect-specific lexical items used in *wh*-question formation (e.g. *inda*, *indalos*, *embu*) or a large number of undisputed phonological differences (not discussed here) — but also, so it seems, with respect to semantico-syntactic restrictions that apply, presenting an interesting arena of comparison. CG in-situ *wh*-phrases, whether dialect-specific or not, quite clearly appear to have different properties from those in SMG, allowing different interpretations in the same environments.

We return to this in section 5, where we offer, if not solutions, at least suggestions how to understand the facts as discussed here, and a little bit beyond). Before going there, however, we would like to first support the data reported in this section and the previous with the results obtained from a grammaticality-judgment questionnaire.

4. The Questionnaire

Validity of the initial observations and intuitions of native speakers as described in sections 2 and 3 was attained through the distribution of a grammaticality-judgment questionnaire.¹¹ This questionnaire was used to explore the possible differences in interpretation which could arise from the different syntactic structures in CG; the results were later compared to SMG (see sections 3 and 5).¹² The questionnaire was set up in order to investigate uses and interpretations of ‘how’ in CG, in particular whether CG *pos* has different semantic and/or syntactic properties from SMG *pos*; it also aimed to identify the properties of CG-specific *indalos*. Differences arising from *wh*-phrases in situ and ex situ in embedded questions were tested as well. It is hypothesized that in-situ *wh*-phrases in embedded questions will be interpreted as the argument of both the matrix and the embedded clause in CG (see e.g. example (27) above), whereas ex-situ *wh*-phrases will be (at least preferably) interpreted as matrix arguments only (see e.g. example (25) above). Any effects of *embu* ‘is(-it)-that’ and referentiality were also tested. Specific items and aspects of design are provided in the appendices.

4.1. Participants

The questionnaire was conducted with thirteen Greek Cypriot native speakers of CG who are permanent residents of Cyprus; only one participant had lived in the UK for 3 years. Since we wanted to test the validity of the initial set of native judgements, we decided to keep the age range constant and thus chose participants aged 20–32 years (M = 25.5, SD = 2.9), balanced for gender (6 female and 7 male). All participants come from an urban background (Nicosia and Larnaca) and none had any linguistic background or other relevant training; the initial informants (see fn. 11) did not participate in the questionnaire.

4.2. Material and Design

The structures and available interpretations of four types of *wh*-questions were tested in the questionnaire, namely those involving the *wh*-arguments *pcos* ‘who-NOM’ and *pcon* ‘who-ACC’ — referred to subsequently and in Appendix B as ‘Who-S(ubject)’ and ‘Who-O(bject)’ — as well as the two *wh*-adjuncts for ‘how’, *pos* (taken over from SMG) and *indalos* (unique to CG). Depending on the syntactic restrictions applying in CG (see section 2), each type of question was distributed evenly across referentiality (R) and non-referentiality (NR), and in-situ, ex-situ, and sentence-medial position of the *wh*-expression were employed. All questions were also distributed along the use or absence of *embu* ‘is(-it)-that’ (which, as mentioned above, will not be reported here any further).

¹¹ Native speaker judgements originally came from the second-named author as well as the reading group participants acknowledged in the title footnote.

¹² Please note that *inda*-questions were not included in the questionnaire, since their status in CG has not been fully established yet (for some discussion, see Grohmann *et al.* 2006, Pavlou, this volume, and Papadopoulou, in progress).

The questionnaire involved two sets of verbs, all checked for frequency. The first set included verbs which could allow for an instrumental reading in *wh*-questions, namely *annio* ‘open’, *katharizo* ‘clean/wipe out’, *kofko* ‘cut’ (as in (20)–(21) above), and the second consisted of three verbs of saying *leo* ‘say’, *sinaferno* ‘talk about’, *murmuro* ‘ramble’ (as in (25)–(27) above). All agents used, male (marked for masculine gender) and female (marked for feminine gender), as well as subjects and objects, are frequently used nouns in CG which were furthermore distributed evenly, along with the verbs, across all conditions (see Appendix A for a full list). This design resulted in 57 sentences which were arranged randomly, so as to avoid any strategies developed by participants (see Appendix B for details).

Specific items in the questionnaire involved question patterns and structures such as those in (28)–(32), that is, *ex-situ* and *in-situ wh*-questions with lexical items that are used in CG without sounding “too Greek” (see also fn. 15 below), including *pos*.

- (28) **Indalos (embu)** aniksen tin kashian o Nikos?
 how EMBU opened the box the Nick
 ‘How did Nick open the box?’
- (29) **Pos (embu)** aniksen tin kashian o Nikos?
 how EMBU opened the box the Nick
 ‘How did Nick open the box?’
- (30) O Nikos aniksen tin kashian **pos**?
 the Nick opened the box how
 ‘Nick opened the box how?’
- (31) **Pcos mitsis (embu)** esinaferen i Maria oti esisen to pulukuin?
 which young-boy EMBU talked-about the Mary that tore the teddy
 ‘Which young boy did Mary say who tore the teddy?’
- (32) I Maria esinaferen oti esisen to pulukui **pcos mitsis**?
 the Mary talked-about that tore the teddy which young-boy
 ‘Mary said which young boy who tore the teddy?’

Items like (33) were also included in the questionnaire, provided in order to clarify the ambiguity observed in (23)–(27) above.

- (33) I Maria esinaferen **pcos mitsis** oti esisen to pulukui?
 the Mary talked-about which young-boy that tore the teddy
 ‘Mary said which young boy who tore the teddy?’

4.3. Procedure

Participants were initially familiarized with obligatory phonological adaptations; among others, double *pp* was used to represent CG /p^h/, a phone that is not part of the SMG inventory. This was considered to be essential, since it helped facilitate for an entirely CG-linguistic environment, avoiding any interaction between SMG and CG.

We leave aside the issue of “artificiality” this choice may be interpreted to cause (see also fn. 13 right below). Note that CG is not orthographically codified, despite recent attempts and a growing body of literature expressed in CG (beyond newspaper articles, there is modern poetry and drama, for example). The SMG writing system is used to write CG, therefore a gap in the representation of double clusters and double consonants is present, as just mentioned.

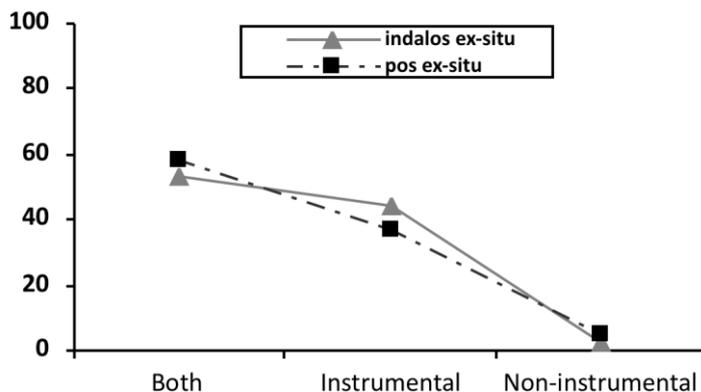
In the absence of a “proper” writing system for CG, the choices we had were using

either Greek or English — arguably neither ideal.¹³ In this sense, the test sentences were written in more or less standard Greek orthography, there were no phonetic clues, and context was not provided (yet a hypothetical interpretation was being elicited from the participants). A three-fold choice was given to the participants with one representing an instrumental reading only, one a manner reading only, and the third indicating both potential interpretations.

4.4. Results

All answers given were coded and analyzed in Microsoft Excel due to a small number of participants, which disallowed for any statistical tests to be run. Initial analysis of the results has shown that *embu* ‘is(-it)-that’, (non-)referentiality (R/NR), and the verbs/nouns used did not have any effect on the results. The *wh*-items *pos* and *indalos* ‘how’ have dissimilar properties, deriving from the fact that they basically are two different lexical items that nevertheless allow for similar readings in the patterns tested. In-situ and ex-situ *pos* seem to employ different strategies in CG, as compared to SMG *pos*. In addition, in-situ and ex-situ Who-O and Who-S questions allow for different readings. Each case is analyzed in detail in the remainder of this section; a more analytical discussion will be presented in section 5.

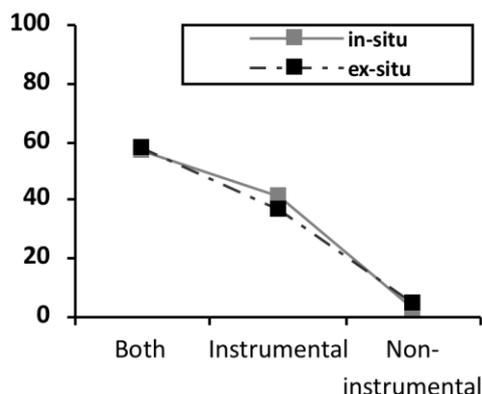
To start with, CG *indalos* does not have the same properties as SMG *pos*, since it can never be left in situ (see section 2 above). As shown in Graph 1, the same pattern is followed with respect to the interpretations allowed with the two *wh*-words; above 50% of the participants allow for instrumental and manner interpretations, and between 37% and 44% allow only for the instrumental reading.



Graph 1: CG *indalos* vs *pos* ‘how’

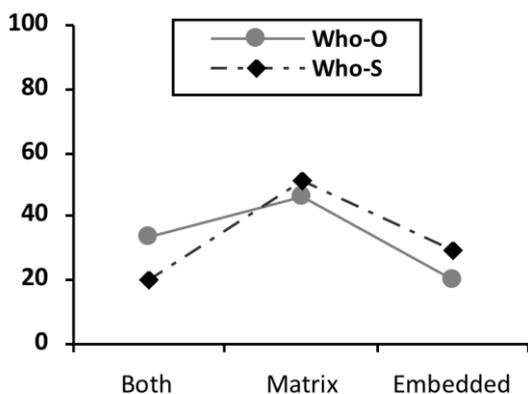
To the extent that *pos* can be used by CG speakers, it can not only be left in situ (as in SMG), but it can also modify the subject as a manner adverb, as opposed to the instrumental-only interpretation in SMG (see section 3). As shown in Graph 2, CG *pos* allows for both interpretations: 56.6% when in-situ and 58% when ex-situ. Accordingly, it is evident that CG *pos* is not affected by its position in the sentence, in contrast to SMG *pos* (again, see sections 2 and 3, but see section 5 for a serious complication of the facts in both languages due to additional evidence reported in Vlachos 2010).

¹³ One might suggest that such research (that is, on linguistic varieties without their own writing systems) better involve auditory presentation of the test sentences, through pre-recorded testing sentences, for example. However, this will not work for the elicitation of quite complex structures and subtle interpretive differences — and especially *wh*-questions — either for either of (at least) two reasons: (i) if recorded with neutral intonation, as would have to be done in order to eschew interpretive effects, the sentences would sound very unnatural, and (ii) if spoken naturally, they would of course give away the intended interpretation(s) immediately.



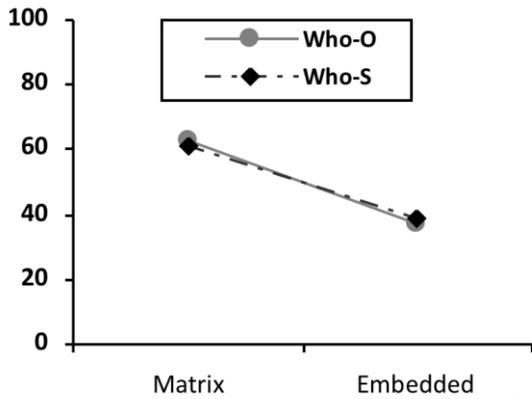
Graph 2: CG pos 'how'

In the presence of a potential ambiguity between a matrix and an embedded reading, interpretation of an ex-situ *wh*-element with the embedded clause is strongly dispreferred, if possible at all, for both Who-O (20%) and Who-S (29%). It becomes clear from Graph 3 that Who-S questions employ a clear dispreference towards the embedded reading, with the choice for both interpretations being lower (20%) than the embedded only (29%). The conclusion we can draw from these results is then: *In complex structures, Who-S correlates most strongly with a matrix-only interpretation, whether ex-situ or in-situ.* In contrast to this, Who-O questions show stronger preference for both interpretations (34%) rather than the embedded only (20%). Still, this allows us to conclude (perhaps a bit weaker): *Who-O correlates most strongly with a matrix-only interpretation.*



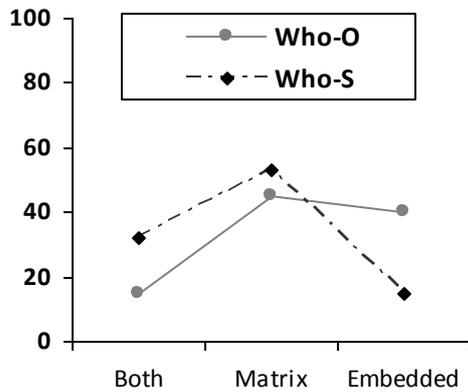
Graph 3: Who-O & Who-S ex-situ

If we accumulate the percentage of the third choice, that is, both to the embedded and the matrix option, as depicted in Graph 4, the same pattern emerges for both types of questions. We can capture this as a firm result as follows: *In the absence of a disambiguating context, wh-ex situ questions in CG complex structures preferably attach a matrix interpretation of the wh-item; an embedded reading is strongly dispreferred.*



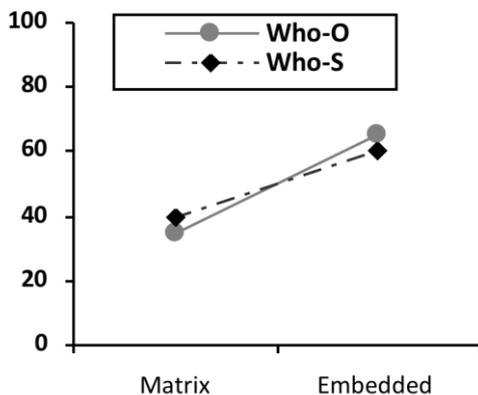
Graph 4: *Who-O/Who-S ex-situ*

In contrast to the above, when the *wh*-word is in situ, the matrix reading is (marginally) possible in CG, unlike SMG where it is clearly ruled out. As represented in Graph 5, there is a clear indifference for matrix readings with *Who-O* questions (15%), whereas for *Who-S*, the embedded interpretation seems to be almost rejected (20%).



Graph 5: *Who-O & Who-S in-situ*

If we break down the “both” options, the same pattern is revealed with a *Who-O* preference for the matrix reading at 35% and a *Who-S* preference at 39.5%, while embedded, the preferences rise to 65% & 60.5%, respectively. This is shown in Graph 6.



Graph 6: *Who-O/Who-S in-situ*

In sum, the quantitative data gathered from the grammaticality-judgment questionnaires administered to 13 CG-native participants confirm the native-speaker intuitions reported in the presentation of the data in sections 2 and 3, by and large. In CG,

the *wh*-items *pos* and *indalos* ‘how’ have dissimilar properties but allow for similar readings, with CG *pos* modifying the subject as a manner adverb either when ex-situ or when left in-situ. In-situ *wh*-expressions in CG (marginally) allow for matrix readings, (i) unlike SMG, and (ii) also in contrast to situations of potential ambiguity between a matrix and an embedded reading, where interpretation of an ex-situ *wh*-element with the embedded clause is strongly dispreferred, if possible at all.

5. Discussion

One result that, we hope, has crystallized throughout the paper so far is that, as discussed in section 2, not every *wh*-item can stay in situ in CG, possibly in contrast to SMG but certainly in line with English, where *how come*, for example, can never appear in situ and where certain *wh*-expressions have been argued to be obligatorily merged “high” (i.e. straight into Spec-C). The same also applies to the CG *wh*-item *indambu*, regardless of whether it is being used argumentally (‘what’) or adverbially (‘why’), and in this respect might differ from English. Certainly, the discussion in the literature concerning *why*, and also *how*, across languages, starting with and inspired by Bromberger (1987), might bear some relevance.

As interesting as it might be, we will not pursue this issue any further other than simply mentioning the fact that certain CG *wh*-expressions can either not stay in situ or never “come” from a lower position to begin with; *(e)mbu* is certainly one of those elements in CG that seem to be obligatorily licensed in the left periphery, whether inserted directly into C (Papadopoulou, in progress) or as the result of a much more complex clefting structure (Grohmann *et al.* 2006); see also Pavlou (this volume) for an overview of several approaches to the shortened variant *mbu* in connection with *inda* (namely, the forms *indambu*, *innambu*, *tambu*, *namu*, and *ambu*, which can all mean ‘what’ or ‘why’). In this sense, we might hold that the ability of a *wh*-expression to appear in situ depends not exclusively on syntactico-semantic licensing options or mechanisms in the grammar, but to a large extent on the lexical properties of a given item.

As a comparative result, a second solid, and arguably the most surprising, difference between CG and SMG *wh*-in situ questions is the availability of a matrix interpretation of an in-situ *wh*-expression in CG that, at least at first glance, appears to occupy a position within an embedded clause — an option which does not exist in SMG. Let us get back to these cases in some more detail by repeating the CG example (27) and providing an additional specimen in (34).

- (27) O Yiannis ipen oti i Maria esinaferen tin aliθkian **se pcon?** [CG]
 the John said that the Mary talked-about the truth to whom
 ‘John said that Mary said the truth to whom?’
- a. Ipen to ston Giorgo.
 said it to-the George
 ‘He said that to George.’
- b. Ipen oti i Maria ipen stin Anna tin aliθkian.
 said that the Mary said to-the Anna the truth
 ‘He said that Mary said the truth to Anna.’
- (34) O Yiannis ipen oti i Maria emourmouran **se pcon?** [CG]
 the John said that the Mary rambled to whom
 ‘John said that Mary rambled on to whom?’
- a. Ipen to ston Giorgo.
 said it to-the George
 ‘He said that to George.’
- b. Ipen oti i Maria emourmouran stin Anna.
 said that the Mary rambled to-the Anna
 ‘He said that Mary rambled on to Anna.’

Recall from the discussion above that SMG does not allow the response in, hence the interpretation construed with, (27a) as well as, by extension, (34a). That is, SMG (*se*) *pjon* ‘(to) whom’ is not able to scope all the way into the matrix, be it by LF-movement or some other licensing operation, whereas CG (*se*) *pcon* seems to be. (35) is a first rough sketch of a possible structural representation (see fn. 10 around the discussion of (24) above for some simplified issues; for us right here, the exact surface positions of subject and verb do not matter):

- (35) [CP **OP** C [o Yiannis ipen ... [vP **OP** (o Yiannis) v [vP _A tv [CP **OP** oti-C
[i Maria esinaferen ... [vP **OP** (i Maria) v [vP tin aliθkian tv **se pcon**]]]]]]]]

The null hypothesis is arguably that *se pcon* originates as the indirect argument of the embedded verb *esinaferen* and then, staying in situ throughout the derivation, somehow takes scope for the (information) interrogative interpretation. Ignoring the matrix clause for the time being, we suggested in (16) above that this “somehow” can be done through unselective binding by an operator **OP** in Spec-C (see Cresti 1998 for discussion, for example) or locally within its immediate domain of interpretation, suggested to be *vP* (see Vlachos 2010 for SMG); the latter we signal through an **OP** in the “edge” of *vP*.¹⁴

If its scope is indeed clause-bound, as argued to hold for SMG (Vlachos 2008, 2010), it should not matter which option we choose: Either the immediate *vP* or the **OP** in the embedded Spec-C might be used to license (*se*) *pcon* in situ — but the result would invariably be an embedded interpretation. This could work for SMG, but not for CG, where a matrix interpretation is acceptable as well. We thus first suggest that something like either (36a) or (36b) could be used for SMG, but not for CG (for simplicity, we use the CG words from (35); replace accordingly with SMG from (26) above, for example):

- (36) a. [o Yiannis ipen ... [CP **OP** oti-C [i Maria esinaferen ...
[vP (i Maria) v [vP tin aliθkian tv **se pcon**]]]]]
b. [o Yiannis ipen ... [CP oti-C [i Maria esinaferen ...
[vP **OP** (i Maria) v [vP tin aliθkian tv **se pcon**]]]]]]

We now return to the matrix clause issue, also relating to (35), and discuss two possible sets of scenarios how matrix interpretation in CG could be integrated into the general picture. The first would require an unselective-binding account for *wh*-in situ and adopt the non-trivial assumption that **OP** in matrix Spec-C may bind the in-situ *wh*-phrase in the embedded clause. Phase-theoretic considerations aside, this assumption is non-trivial in that one would have to claim — and ideally, support with additional data — that CG *wh*-in situ differs from SMG in not being restricted to a single clause boundary. We currently have no such additional data, and neither do we have any reason to believe that CG would indeed differ from SMG in this respect. In this case, the **OP** in matrix Spec-C in (35) would be the licensing operator. So instead of (36a) for SMG, we would be dealing with (37a) for CG; for the *vP*-licensing account, it would be (37b).

- (37) a. [CP **OP** C [o Yiannis ipen ... [i Maria esinaferen ...
[vP (i Maria) v [vP tin aliθkian tv **se pcon**]]]]]
b. [CP [o Yiannis ipen ... [vP **OP** v [i Maria esinaferen ...

¹⁴ In case it has not transpired yet, our goal here is not to come up with the best possible analysis for, or even a novel account of, licensing *wh*-in situ — be it for Greek or more generally. Rather, we would like to try to make sense of the structures and interpretations our study has uncovered. We thereby might cut some corners and possibly avoid further discussions in a nonchalant manner by somewhat simplifying or glancing over details, but we hope that the tools and assumptions we employ here are transparent enough, yet interesting and relevant.

[_{vP} (i Maria) v [_{vP} tin aliθkian t_v **se pcon**]]]]

The unselective-binding account from matrix Spec-C in (37a) can only be made to work if CG *wh*-in situ is *not* clause-bound. The same can be said for the local-licensing account in (37b), except that in this case, it would not even be “local” anymore. Vlachos’ (2010) proposal that *wh*-in situ expressions are licensed locally, within their immediate vP, is exactly that: presupposing that their interpretation is clause-bound and evoking the local, i.e. immediately dominating, vP. Neither is given in (37), so we discard this first set of scenarios flat out. We thus need a more satisfactory account.

As an alternative, we capitalize on the additional position in (35) marked, as in (24) above, _A. The rough story of the second scenario is that the two interpretations arise from an ambiguous lexical choice: *ipe* ‘said’ used monotonically vs. ditransitively.¹⁵ That is, looks are deceiving and *se pcon* in (27) is not in situ in the embedded clause after all but rather in the matrix clause, roughly in the position of _A.

Implementing this idea, we could revise our structure(s) for CG (27) and account for the availability of a matrix interpretation of the apparently embedded in-situ *wh*-item through structural ambiguity:

- (38) a. [**OP** o Yiannis ipen ... [_{vP} **OP** (o Yiannis) v [_{vP} **se pcon** t_v
[_{CP} oti i Maria esinaferen tin aliθkian]]]]
b. [o Yiannis ipen ... [_{CP} **OP** oti-C [i Maria esinaferen ...
[_{vP} **OP** (i Maria) v [_{vP} tin aliθkian t_v **se pcon**]]]]]]

Here *se pcon* is either generated as the indirect object of the matrix verb *ipen* ‘said’ (for example, in Spec-V, as in (38a)) or originates in the embedded clause, as the indirect object of *esinaferen* ‘talked’ (as in (38b)). The “good news” is that these structures again allow both the unselective-binding or the local-licensing accounts of *wh*-in situ, as signaled by the positions for OP, under which each instance of (*se*) *pcon* would be licensed (immediately) within its respective clause.

Note two things first, however: (i) the (external) merge position of *se pcon* would be different in the two cases, as illustrated in (38), even though they arguably play identical roles as indirect objects; (ii) if *se pcon* were merged as an argument of the matrix verb to yield the matrix reading, it would not come out as such in an in-situ linearized string — it is not in the “final” position in which it is pronounced. Perhaps neither objection is terribly worrying, in which case we leave the choice to the reader (see also fn. 14 above). After all, the finer structure of vP might need revising anyway, and the jury is still out on how linearization really works and when it applies in the derivation.

Whichever way to go, it becomes clear that under anyone’s take on scope and interpretation, a matrix reading of (*se*) *pcon* in cases like (27) and others requires that at some point in the derivation, (*se*) *pcon* passes through the matrix clause. If Vlachos’ (2008, 2010) discussion of clause-boundedness of SMG *wh*-in situ extends to CG, this can only mean that it must have started out there. That is to say, *se pcon* must originate in the matrix clause, roughly as in (38a), otherwise it cannot be construed with matrix

¹⁵ This is, of course, why Vlachos (2008) chose the SMG verbs *anakinose* ‘announced’ and *apokalipse* ‘revealed’. However, CG purportedly does not make this subtle distinction, so we opted for using the most natural CG verb of saying, *ipe*, the past tense of *leo* ‘say’ (see also Appendix A for a list of verbs used). If we had used Vlachos’ verbs, the respondents would invariably have perceived an SMG-influenced tone in the test sentences, unnatural for CG, and might perhaps have responded differently. Note that we used several different verbs, however, each one alternating in matrix and embedded contexts, without significant effects.

This issue clearly reflects the difficulties not only for investigating varieties without a writing system through a written questionnaire (see section 4.3 and fn. 13 above), but also the sensitive task of exploring a “low-prestige” variety (CG), trying not to find or create interference from the “high-prestige” variety (SMG).

interpretation. Leaving aside for now the exact licensing mechanism(s) of *wh*-in situ more generally (i.e. whether it is through Spec-C, Spec-v, or some other manner), this means that *se pcon* either starts out as an argument, as in (38a), and something else needs to be said on linearizing it properly — or it is right-adjoined from the start (to matrix VP/vP), again leading to non-trivial consequences. We will explore this option for *pos* presently.

First, however, we briefly address those complex interrogative structures with *wh*-ex situ, for which we observed a clear difference between CG and SMG: The embedded interpretation of the *wh*-item is strongly dispreferred. This was the case for (25), repeated here:

- (25) **Se pcon (*embu*)** esinaferen o Yiannis oti i Maria ipen tin alithkian? [CG]
 to whom EMBU talked-about the John that the Mary said the truth
 ‘To whom did John say that Mary said the truth?’
- a. Ipen to ston Giorgo.
 said it to-the George
 ‘He said it to George.’
- b. # Ipen oti i Maria ipen stin Anna tin aliθkian.
 said that the Maria said to-the Anna the truth
 ‘He said that Mary said the truth to Anna.’

In the absence of additional evidence, we assume (25) to be the ex-situ version of (27), minus the optional *embu* (discussed in section 2) and with the verbs reversed (but see the brief comment in fn. 13 that the matrix vs. embedded appearance of the chosen verb had no significant effect on interpretation).

If so, a version of (38) should underlie the derivation of (25) as well, that is, in theory *se pcon* should be generated either in the matrix clause (39a) or in the embedded clause (39b):¹⁶

- (39) a. [**se pcon** *embu* esinaferen ... [_{VP} o Yiannis v [_{VP} (**se pcon**) t_v]
 [_{CP} oti i Maria ipen tin aliθkian]]]]
- b. [**se pcon** *embu* esinaferen o Yiannis [_{CP} (**se pcon**) oti-C
 [i Maria ipen [_{VP} tin aliθkian t_v (**se pcon**)]]]]]]

Again, these are possibly the underlying derivations for SMG (see (23) in section 3, discussed in Vlachos 2008, 2010) for which, again, the corresponding lexical items from SMG should be inserted. But for CG, (39b), at least, seems to be inappropriate, since it would predict that the moved *wh*-expression (*se*) *pcon* should be able to reconstruct and yield the embedded reading — which is not available.

To be honest, we do *not* have an interesting explanation for this state of affairs, if any at all. One factor we assumed would not seem to play a role is the choice of verb. As mentioned before (e.g., fn. 15), the three verbs of saying we used, namely *leo* ‘say’, *sinaferno* ‘talk about’, and *murmuro* ‘ramble’ (see section 4.2), did not exhibit any effects on the participants’ responses. Note, first of all, that the CG verb *sinaferno* does not exist in SMG (Babinotis 2008). It is a verb derived from the Ancient Greek *sinanafero*, used when talking about someone who is not present (Giagoulis 2009:455). We gather from our informants that it is nowadays used synonymously with *leo*, which is why we consistently translated it as ‘say’ in the data presented here. The test sentences contained one of these three verbs in the matrix and the embedded clause, but never the same verb twice in a given sentence. We thus deemed it unlikely that the verb form *esinaferen* ‘talked about’ in

¹⁶ Here we signal the original, externally merged copy of *se pcon*, as well as the purported intermediate copy in (39b), with boldfaced parentheses and gloss over structural details irrelevant at this point (labels of projections, position of subjects, and other aspects of the derivational history; see also fn. 10 above).

examples like (25) had a particular effect, since in other sentences it appeared in the embedded clause, yet the embedded interpretation was not construed. However, since (40) is such an example, our initial assumption might not be so innocent and straightforward after all. We admit that in this instance, *esinaferen* is best translated as ‘talked about’. In fact, since its argument structure seems saturated by the clitic *ton*, it is impossible to construe an embedded interpretation of the *wh*-moved PP *se pcon*.

- (40) **Se pcon (*embu*)** ipen i Maria o Nikos esinaferen ton? [CG]
 to whom EMBU said the Maria that the Nick talked-about him-CL
 ‘To whom did Mary say that Nick talked about him?’

Unless we are overlooking some crucial aspect of CG grammar, the facts seem to turn out the way described here. Syntactically, (25) and similar data might suggest that CG does not allow long (*wh*-) movement, which would be wrong; hence, we will not pursue this option. Neither will we pursue an oft-heard assessment of speakers, something to the effect of: “Cypriots don’t like to use complicated sentences.” We thus cannot offer a decent explanation for this aspect.

A final intended result of our study was to show a discrepancy between CG and SMG as regards the availability of instrumental and manner readings with *pos*-in situ. This would have been the most puzzling difference, primarily for theoretical reasons, as the following discussion will bring to light. Alas, things are never that simple, so let’s roll this up from the beginning. This final part of our discussion leads us then to the purported difference between SMG (21) and CG (22), the latter of which repeated here for convenience, where, in contrast to CG, the b-response was reported to be infelicitous for SMG in Vlachos (2008):

- (22) O Nikos aniksen tin porta **pos**? [CG]
 the Nick opened the door how
 ‘Nick opened the door how?’
 a. Me to kliði.
 with the key
 ‘With the key.’
 b. Nevriasmenos.
 angry-NOM
 ‘With anger.’

As mentioned in the title footnote, Vlachos (2008), on which we based our original investigation, was subsequently revised and appeared as Vlachos (2010). The revisions include some of the data reported earlier, and the published version differs in crucial respects as regards both the analysis and the treatment of a number of data. One of these concerns cases like (21) in SMG. Vlachos (2008) reports that the predicate adjective *nevriasmenos* ‘angry-NOM’ would in this case be infelicitous, unlike CG, as shown in (22b). Three notes are in order, however, leading to another possibly extended discussion.

First, somewhat surprisingly perhaps, Vlachos employed the adjective *nevriasmenos* as opposed to the adverb *nevriasmena*. And indeed, as he reports in his published work, the adverb is acceptable for SMG speakers. The “updated” (41) is taken from Vlachos (2010).

- (40) O Nikos anikse tin porta **pos**? [SMG]
 the Nick opened the door how
 ‘Nick opened the door how?’
 a. Me to kliði.
 with the key
 ‘With the key.’
 b. Nevriasmena.
 angrily

‘Angrily.’

It is thus possible, even in SMG, that *pos*-in situ may have a subject-oriented manner interpretation, as in CG.

Second, Vlachos notes in this context: “Although for some Greek native speakers the subject-related reading of the *wh*-in-situ adverb does not immediately derive” (Vlachos 2001:fn.3). There is thus some additional variability which should be taken into account in further studies of this phenomenon.

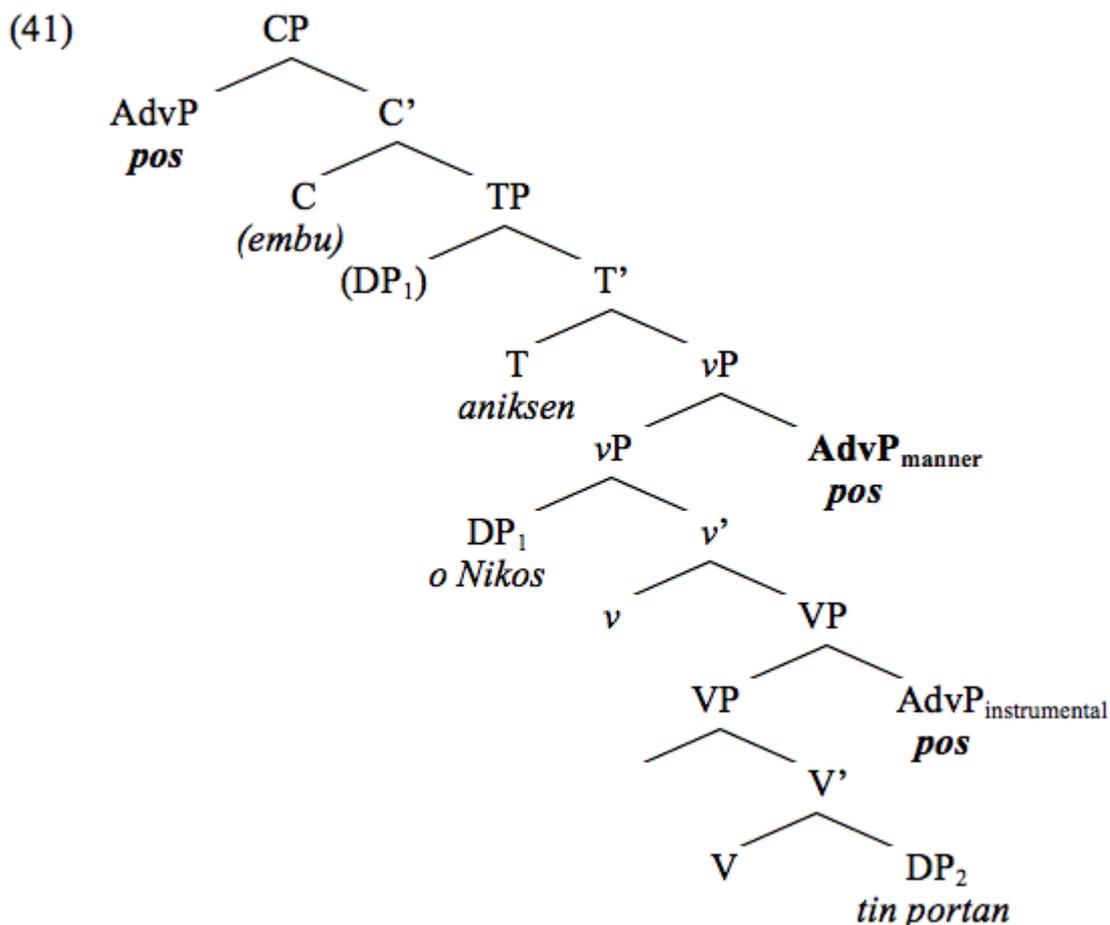
Third, unlike the “original” discussion by Vlachos (2008), the pilot of our questionnaire contained the adverb *nevriasmēna* instead of the adjective *nevriasmēnos*. However, the reaction of native speakers was that the adverb sounded “too Greek” (SMG-like), and that they preferred *nevriasmēnos*.¹⁷ Having opened one can of worms too many already, we will not venture into a monologue on the CG use of adjectives vs. adverbs, or some deeper grammatical variation in this large area between the two varieties, but we at least take the speakers’ intuitions seriously that in this context, they prefer the adjective, and this adjective seems to be less preferred in the same context by SMG speakers, as reported in Vlachos (2008) and several other speakers of SMG we consulted afterwards. We also elicited five additional CG judgements on *pos*-ex and -in situ post-hoc, with the adjective, and all five speakers went for both interpretations in both contexts, thereby confirming that (20) and (22) do indeed hold for CG. Moreover, when asked how they interpreted *nevriasmēnos*, all five responded (again, in both contexts): “Nick opened the door with anger.”

Thus, while, in light of Vlachos (2010), our results concerning *pos*-in situ may not appear as strong as they did compared to Vlachos (2008), the situation for CG, at least, seems clear: The in-situ and the ex-situ use of *pos* ‘how’ allows a subject-oriented manner reading as well as an instrumental interpretation. How significant this result is with respect to SMG is another matter.

On the analytical side, however, comparing the discrepancy between SMG (21) and CG (22), with the adjective *nevriasmēnos* ‘angry’ as intended “manner reading” and the PP *me to kliði* ‘with the key’ as “instrumental reading” for the *wh*-adverbial *pos* ‘how’, we would like to suggest the difference between CG and SMG to lie in one (or both) of two factors: (i) CG has the manner-adverbial right-adjoined to *vP*, scoping immediately over the thematic subject, while SMG only allows right-adjunction of *pos* to *VP* and (ii) only in CG can the manner adverbial stay in situ, while in SMG *pos* has to *wh*-move obligatorily. The structural option is depicted in (41):

(41)

¹⁷ This assessment came from all five speakers pre-tested. Of course, we refer to our fn. 15 above once more.



Perhaps the structure in (41) can be improved in the near future, but without further testing, neither option seems particularly appealing at the moment. Once again, we have to defer a final analysis to a better understanding of Greek (dialectal) syntax of which we are, admittedly, not specialists. But we would like to submit the conjecture that unless a working solution can be found, matters would be rather strange if the clause structure and derivational histories for these constructions were by and large the same for CG and SMG, as we tentatively assume here. This is not to say that the clause structure or syntactic derivation may not differ at all, quite the opposite. Since at least Terzi (1999a, 1999b), the idea has been pursued that the differences in clitic placement (SMG proclisis vs. CG enclisis in many identical syntactic environments) might lie in a different landing site of the verb, which in CG would move to a higher position than in SMG. For extensive, and more recent, discussion on clitic placement, see especially Mavrogiorgos (2009). So some structural differences might be present (CG might employ a different functional head), going hand in hand with derivational differences (that this head attracts the verb only in CG), but with respect to *wh*-in situ structures or finer and more subtle interpretation differences, we just lack the relevant data at this point to warrant such a hypothesis.

This leads us to a final postscript on earlier analytical forays we ventured into. We simply do not have enough facts to say with some certainty that the structural or derivational properties of the two varieties differ in significant ways for the cases at hand — be it *pos*-in situ and differing interpretation construals, be it *pcon*-ex situ and the absence of an embedded reading in CG, or be it *pcon*-in situ and the availability of either construal in CG. As such, we tried to restrict ourselves in this paper to discuss some initial observations, then also corroborated quantitatively, about the “facts” as we presented them here, added by several digressions on various analytical paths one could tread on towards an understanding of why CG and SMG seem to diverge the ways they

seem to. What we will not do is pursue yet another route of “explanation” — one that might build on the above-mentioned oft-heard assessment of Greek Cypriots that their language would not allow “complicated” sentences (see also fn. 2 in above). A cynic might take this to the creole route of CG phylogeny, but we will resist such temptation. A more constructive reading of this assessment (or whatever there is to it) might be a processing account, which we will likewise ignore for now. Absence of an embedded interpretation of *pcon*-in and *ex situ* in complex structures could be construed as a “preference” to process top-down and stop interpreting once a first possible reading has been found — in either case the matrix position, whether “real” or not. Needless to say, while this might help account for why shorter movement should generally be preferred, it would make for a very weak case for a(ny) derivational approach.

Note finally that stating anything valid about the grammar of CG is notoriously difficult: CG is considered by many speakers not to be a *bona fide* variety in the first place, carrying “low prestige”; in addition, and related to this point, Greek Cypriots tend to perceive CG as “inferior” in some way and consequently look down upon their own language; Papapavlou (1998), for example, investigated speakers’ attitudes by marking 12 traits such as kindness, intelligence, sincerity, dependability, and sense of humor carried by CG versus SMG.¹⁸ As a result, it is tremendously difficult to extract stable judgments shared by a majority of speakers.

At the same time, as we have already mentioned (and again, related to both previous points), there is no codified, official grammar of CG — although, and this makes us hopeful for future research, several such enterprises are currently on their way, such as the Kykkos Monastery’s *Thesaurus Linguae Cypriae Graecae* project (<http://www.thisavros.com>). In addition, the body of formal research on CG morphosyntax is constantly growing (starting with work on clitics by Agouraki 1997 and Terzi 1999a, 1999b, but expanding more and more, such as Grohmann *et al.* 2006, among several others). Also, it can only be hoped that research on CG acquisition (for the longest time restricted, more or less to, Petinou & Terzi 2002, but currently expanded by Neokleous, in progress, also on the acquisition of clitics and Papadopoulou, in progress, on *wh*-related issues), in particular the systematic investigations into child language development in typically developing and language-impaired children carried out by the Cyprus Acquisition Team (<http://www.research.biolinguistics.eu/CAT>) will eventually bear fruit as well. The latter is going to be done within the newly funded *Gen-CHILD* research project (see Grohmann 2010, upcoming for overviews).

Note also that the influence of SMG, and the role it plays in daily life and society in (Greek-speaking) Cyprus, surely needs to be taken into account, which we have not. This is, of course, one of the perpetual problems with “diglossia” (see Papapavlou & Pavlou 1998, but also Karyolemou 2006) — and, for most people, gives rise to the question: Where, when, and how does a variety become a grammar? (As already mentioned in section 3, this is not so for most, if not all, generative linguists; see, among many others, Kayne (2000) for extensive discussion and argumentation from a “micro-parametric” perspective applied to the myriad varieties of the Romance language family.) In this context, one should also be more careful with one’s research of “the” CG variety/idiolect/dialect/language — factors such as “bleaching” from SMG and others, as well as (possibly, but not necessarily, geographical) variation within CG itself, may further

¹⁸ As a way of “revising” the perhaps bleak tone of fn. 2 above, we would like to point out that, while there certainly are Greek Cypriots that reject CG as a proper language or at least look down upon it, there is also an increasing number that feels the exact opposite. See, among others, Moschonas (2002), who also cites the above-mentioned Papapavlou (1998) for positive attitudes as well (see also Papapavlou 2001), Gardner-Chloros *et al.* (2006), and Karyolemou (2006). For a recent study, see especially Leivada *et al.* (2009), and references cited there, regarding attitudes towards CG and linguistic change in Cyprus; the authors report on their survey in which 51 out of 80 participants expressed a desire for the recognition of CG as the official language of the Republic of Cyprus.

blur the issue. Solid sociolinguistic research, such as Pappas (in press) on variability in CG clitic placement, is noteworthy in this respect.

And lastly, even if these points can be dealt with, the design of our data gathering may be argued to leave room for improvement: the number of speakers tested (13 respondents is not very representative), the time it takes to fully fill out the questionnaire (57 sentences take a long time to process), or the presentation of the test sentences (Greek orthography for oral CG, absence of phonetic clues, lack of context with a request for interpretation, and so on), possibly among other factors as well.

Such concerns notwithstanding, we would like to close our report on a more positive note. This study has shown that there are serious grammatical differences between “low” CG and “high” SMG which can be investigated formally, even in the presence of obstacles.

In addition, a growing body of work is currently being devoted to language development, specifically to the first language acquisition of CG by typically developing children (as well as language-impaired children). Activities of the above-mentioned Cyprus Acquisition Team, a research group recently initiated by the first-named author and now officially funded (cf. Grohmann 2010, upcoming, to appear), have already started looking into the acquisition of *wh*-questions, and the research about to be completed by the second-named author deals with very similar issues (Papadopoulou, in progress). In the future, we will develop a modified testing tool to determine the onset of interpretations such as those discussed here, or even the availability of in-situ information questions, with young children.

Appendix A: List of Verbs, Nouns, and Agents

Verbs (matrix)	<i>annio</i> 'open' <i>katharizo</i> 'clean' <i>kofko</i> 'cut' <i>leo</i> 'say' <i>sinaferno</i> 'talk about' <i>murmuro</i> 'ramble'
Verbs (embedded)	<i>pao</i> 'go' <i>derno</i> 'hit' <i>shizo</i> 'tear'
Nouns	<i>vazanin</i> 'aubergine' <i>kashia</i> 'box' <i>aftokinito</i> 'car' <i>vurna</i> 'sink' <i>alithkia</i> 'truth' <i>peripatos</i> 'walk' <i>pulukuin</i> 'teddy'
Agents	<i>Yiannis</i> 'John' <i>Nikos</i> 'Nick' <i>Anna</i> 'Anna' <i>Maria</i> 'Mary'
Agents (controls)	<i>mitsis</i> 'young boy' <i>mastros</i> 'boss'

Appendix B: Distribution of Experimental Conditions

item no.	<i>wh</i> -word	R/NR	in-situ	first V (matrix)	second V (embedded)	<i>embu</i>	first N (matrix)	second N (embedded)
1	<i>who</i> -O	R	-	V4	V6	+	M1	F2
2	<i>who</i> -S	N R	-	V6	X3	-		
3	<i>inda</i> <i>los</i>	N/ A	-	V3	N/A	-	M2	
4	<i>who</i> -O	R	+	V6	V5	-	F2	M1
5	<i>who</i> -S	N R	-	V4	X1	+		
6	<i>who</i> -S	R	+	V5	X2	-	F1	
7	<i>who</i> -O	N R	-	V5	V6	-	M2	F1
8	<i>pos</i>	N/ A	-	V1	N/A	+	F2	
9	<i>who</i> -S	R	±	V4	X1	-		
0	<i>who</i> -O	N R	+	V4	V6	-	F1	M2
1	<i>who</i> -S	R	-	V6	X3	+	M2	
1	<i>pos</i>	N/ A	+	V3	N/A	-	F2	
1	<i>who</i> -S	R	-	V5	X2	-		
1	<i>who</i> -O	N R	-	V6	V5	+	F1	M2

1 5	<i>who</i> -S	R	+	V4	X1	-	M2	
1 6	<i>who</i> -O	R	-	V4	V6	-	F2	M1
1 7	<i>inda</i> <i>los</i>	N/ A	-	V1	N/A	+	M1	
1 8	<i>who</i> -S	N R	±	V4	X1	-	F1	
1 9	<i>who</i> -S	N R	-	V5	X2	-	M2	
2 0	<i>who</i> -o	R	-	V6	V4	+	F2	M1
2 1	<i>pos</i>	N/ A	-	V1	N/A	-	M1	
2 2	<i>who</i> -O	R	+	V5	V6	-	F1	M2
2 3	<i>who</i> -S	N R	-	V6	X3	+		
2 4	<i>inda</i> <i>los</i>	N/ A	-	V2	N/A	-	F2	
2 5	<i>who</i> -S	N R	+	V6	X3	-		
2 6	<i>pos</i>	N/ A	-	V3	N/A	-	F1	
2 7	<i>who</i> -O	N R	-	V6	V5	-	M2	F1
2 8	<i>who</i> -S	N R	±	V5	X2	-		
2 9	<i>who</i> -S	R	-	V4	X1	+	M1	
3 0	<i>pos</i>	N/ A	+	V2	N/A	-	F1	
3 1	<i>who</i> -S	R	-	V4	X1	-	M2	
3 2	<i>who</i> -O	N R	-	V5	V6	+	F2	M1
3 3	<i>pos</i>	N/ A	-	V2	N/A	-	M1	
3 4	<i>who</i> -O	N R	+	V5	V4	-	F1	M2
3 5	<i>inda</i> <i>los</i>	N/ A	-	V3	N/A	+	M2	
3 6	<i>who</i> -O	R	-	V5	V4	-	F2	M1
3 7	<i>who</i> -S	R	+	V6	X3	-	M1	
3 8	<i>who</i> -O	R	-	V5	V4	+	F1	M2
3 9	<i>pos</i>	N/ A	+	V1	N/A	-	M2	
4 0	<i>who</i> -S	R	±	V5	X2	-	F2	
4 1	<i>who</i> -S	N R	-	V5	X2	+	M1	
4 2	<i>who</i> -O	R	+	V4	V5	-	F1	M2
4 3	<i>who</i> -S	N R	-	V4	X1	-		
4 4	<i>pos</i>	N/ A	-	V2	N/A	+	F2	
4 5	<i>who</i> -S	R	+	V5	X2	-		
4 6	<i>who</i> -O	N R	-	V4	V5	-	F1	M2

4 7	<i>who</i> -S	R	-	V5	X2	+		
4 8	<i>inda</i> <i>los</i>	N/ A	-	V1	N/A	-	F2	
4 9	<i>who</i> -S	R	±	V6	X3	-		
5 0	<i>who</i> -O	N R	-	V4	V5	+	F1	M2
5 1	<i>pos</i>	N/ A	-	V3	N/A	-	M2	
5 2	<i>who</i> -S	R	-	V6	X3	-	F2	
5 3	<i>who</i> -O	N R	+	V6	V4	-	M1	F2
5 4	<i>inda</i> <i>los</i>	N/ A	-	V2	N/A	+	F1	
5 5	<i>who</i> -S	N R	+	V4	X1	-		
5 6	<i>who</i> -O	R	-	V6	V4	-	F2	M1
5 7	<i>who</i> -S	N R	±	V6	X3	-	M1	

References

- Adli A. (2006). French *Wh*-in-situ Questions and Syntactic Optionality: Evidence from Three Data Types. *Zeitschrift für Sprachwissenschaft* 25: 163-203.
- Agouraki Y. (1997). On the Enclisis/Proclisis Alternation. In G. Drachman, A. Malikouti-Drachman, J. Fykias and C. Klidi (eds.) *Proceedings of the 2nd International Conference on Greek Linguistics*, vol. II. Salzburg: University of Salzburg, 393-404.
- Alexiadou A. & Anagnostopoulou E. (1998). Parametrizing Agr: Word Order, V-Movement and EPP-Checking. *Natural Language and Linguistic Theory* 16: 491-539.
- Babinotis G. (2008). *Λεξικό της Νέας Ελληνικής Γλώσσας* (Dictionary of Modern Greek), 3rd edn. Athens: Lexicology Centre.
- Belletti A. (2004). Aspects of the Low IP Area. In L. Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2. New York: Oxford University Press, 16-51.
- Bolinger D. L. (1978). Asking More than One Thing at a Time. In H. Hiz (ed.), *Questions*. Dordrecht: D. Reidel, 107-150.
- Bošković Ž. (1998). LF Movement and the Minimalist Program. *Proceedings of NELS* 28: 43-57.
- Bromberger S. (1987). What We Don't Know When We Don't Know Why. In N. Rescher (ed.), *Scientific Inquiry in Philosophical Perspective*. Lanham, MD: The University Press of America, 75-104. [Reprinted in Bromberger, S. 1992. *On What We Know We Don't Know: Explanation, Theory Linguistics, and How Questions Shape Them*. Chicago: University of Chicago Press and Stanford, CA: Center for Study of Language and Information, 145-169.]
- Chang (1997). *Wh-in-situ Phenomena in French*. University of British Columbia master thesis, Vancouver.
- Cheng L. L.-S. & Rooryck J. (2002). *Types of Wh-in-situ*. Unpublished MS. Universiteit Leiden.
- Chomsky N. 2000. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press, 89-155.
- Collins C. (1991). *Why and How Come*. *More Papers on Wh-Movement — MIT Working Papers in Linguistics* 15: 31-45.
- Contossopoulos N. (1983-1984). La Grèce du $\tau\iota$ [ti] et la Grèce du $\epsilon\iota\nu\tau\alpha$ [inta]. *Glossologia* 2-3: 149-162.
- Cresti D. (1998). Some Considerations on *Wh*-Decomposition and Unselective Binding. In G. Katz, S.-S. Kim and H. Winhart (eds.), *Sprachtheoretische Grundlagen für Computerlinguistik: Arbeitspapiere des Sonderforschungsbereichs 340*. Tübingen: Universität Tübingen, SFB 340, 155-183.
- Fotiou C. (2009). Focusing Strategies in Cypriot Greek. In K. K. Grohmann and P. Panagiotidis (eds.) *Selected Papers from the 2006 Cyprus Syntaxfest*. Newcastle-upon-Tyne: Cambridge Scholars

- Press, 63-91.
- Gardner-Chloros P., McEntee-Atalianis L. & Finnis K. (2006). Language Attitudes and Use in a Transplanted Setting: Greek Cypriots in London. *International Journal of Multilingualism* 2: 52-80.
- Giagoulis K. G. (2009). *Θησαυρός Κυπριακής Διαλέκτου* (Thesaurus of the Cypriot Greek Dialect). Nicosia: Theopress.
- Ginzburg J. & Sag I. (2000). *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, CA: CSLI.
- Grohmann K. K. (2003). *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.
- Grohmann K. K. (2010). Explorations in the Acquisition of Cypriot Greek: A Preliminary View from the Gen-CHILD Project. Paper presented in the *CLT Seminar Series*, Centre de Lingüística Teòrica, Universitat Autònoma de Barcelona. (March 23, 2010).
- Grohmann, K. K. (Upcoming). The Dialect Design: Socio-Syntax of Development and the Grammar of Cypriot Greek. Invited presentation at *The Language Design* workshop, Université du Québec à Montréal. (May 28–30, 2010).
- Grohmann, K. K. (2010 in press). Anti-Locality: Too-Close Relations in Grammar. In C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press.
- Grohmann K. K. (To appear). Issues in and for the Acquisition of Cypriot Greek. In E. Rinke and T. Kupisch (eds.), *The Development of Grammar: Language Acquisition and Diachronic Change*. Amsterdam: John Benjamins.
- Grohmann K. K., Panagiotidis P. & Tsiplakou S. (2006). Properties of *Wh*-Question Formation in Cypriot Greek. In M. Janse, B. D. Joseph and A. Ralli (eds.), *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory (Mytilene, Greece: 30 September–3 October 2004)*. Patras: University of Patras, 83-98.
- Grohmann K. K. & Papadopoulou E. (Forthcoming). Question(able) Issues in Cypriot Greek. *Linguistic Analysis* 37 — Special Issue: 'On the Optionality of *Wh*-Movement' (guest-edited by Anna Roussou and Christos Vlachos).
- Gryllia S. & Lekakou M. (2006). Syntax Reflects Information Structure: Evidence from Cypriot Greek. Paper presented at the *27th Annual Meeting of the Department of Linguistics*, Aristotle University of Thessaloniki (May 6-7, 2006).
- Heim I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. University of Massachusetts doctoral dissertation, Amherst.
- Holton D., Mackridge P. & Philippaki-Warbuton I. (1997). *Greek: A Comprehensive Grammar of the Modern Language*. London: Routledge.
- Huang C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. MIT: PhD dissertation, Cambridge.
- Ioannidou A., Mavrada S., Pavlou N. & Sophokleous E. (2009). The Awareness of Being Aware: The Case of Cyprus. Unpublished MS. University of Cyprus, Nicosia.
- Kamp H. (1981). A Theory of Truth and Semantic Representation. In J. Groenendijk, T. Janssen and M. Stokhof (eds.), *Formal Methods in the Study of Language*. Amsterdam: Universiteit van Amsterdam, 277-320.
- Karyolemou M. (2006). Reproduction and Innovation of Communicative Patterns in a Former-"Diglossic" Community. In R. Muhr (ed.) *Reproduktionen und Innovationen in Sprache und Kommunikation verschiedener Sprachkulturen/Reproduction and Innovation in Language and Communication in Different Language Cultures*. Frankfurt am Main: Peter Lang, 39-56.
- Kayne R. (2000). *Parameters and Universals*. New York: Oxford University Press.
- Ko H. (2005). Syntax of *Why*-in-situ: Merge into [Spec, CP] in the Overt Syntax. *Natural Language and Linguistic Theory* 23: 867-916.
- Leivada E., Hadjiefthymiou P. & Mavroudi P. (2009). Attitudes towards Low Varieties and Linguistic Change: The Case of Cypriot Greek. Unpublished MS., University of Cyprus, Nicosia.
- Mathieu E. (2004). The Mapping of Form and Interpretation: The Case of Optional *Wh*-Movement in French. *Lingua* 114: 1090-1132.
- Mavrogiorgos M. (2009). *Proclisis and Enclisis in Greek*. University of Cambridge doctoral dissertation.
- Moschonas S. (2002). Κοινή γλώσσα και διάλεκτος: Το ζήτημα της «γλωσσικής διμορφίας» στην Κύπρο (Common Language and Dialect: The Case of 'Linguistic Diglossia' in Cyprus). *Nea Estia* 151: 898-928.
- Munaro N. & Pollock J.-Y. (2005). *Qu' est-ce que (qu)-est-ce que?* A Case Study in Comparative

- Romance Interrogative Syntax. In G. Cinque and R. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, 542-606.
- Neokleous T. (in progress). *L1 Acquisition of Cypriot Greek Pronominal Clitics*. University of Cambridge, PhD dissertation.
- Newton B. (1972). *Cypriot Greek: Its Phonology and Inflections*. The Hague: Mouton.
- Ochi M. (2004). *How Come* and Other Adjunct *Wh*-Phrases: A Cross-Linguistic Perspective. *Language and Linguistics* 5: 29-57.
- Papadopoulou, E. (in progress). *Acquisition of Wh-Questions: Evidence from Cypriot Greek*. University of Essex, PhD dissertation, Colchester.
- Papagelou R. (2001). *Το Κυπριακό Ιδίωμα: Μέγα Κυπρο-Ελληνο-Αγγλικό Λεξικό* (The Cypriot Dialect: A Large Cypriot-Greek-English Dictionary). Athens: Iolkos.
- Papapavlou A. N. (1998). Attitudes toward the Greek Cypriot Dialect: Sociolinguistic Implications. *International Journal of the Sociology of Language* 134: 15-28.
- Papapavlou A. N. (2001). Mind Your Speech: Language Attitudes in Cyprus. *Journal of Multilingual and Multicultural Development* 22: 491-501.
- Papapavlou A. N. & Pavlou P. (1998). A Review of the Sociolinguistic Aspects of the Greek Cypriot Dialect. *Journal of Multilingual and Multicultural Development* 19: 212-220.
- Pappas P. A. (in press). Object Clitic Placement in the History of Cypriot Greek. In B. Heselwood & C. Upton (eds.), *Proceedings of Methods XIII*. Frankfurt am Main: Peter Lang.
- Pavlou N. (2009). Flavours of *mbu* in Cypriot Greek. Paper presented at the 9th International Conference of Greek Linguistics, University of Chicago (October 29-31, 2009).
- Pavlou N. (2010). *Mbu!* On *Wh*-Objects and True Adjuncts in Cypriot Greek. Unpublished MS., University of Cyprus, Nicosia.
- Pesetsky, D. 1987. *Wh*-in-situ: Movement and Unselective Binding. In E. Reuland and A. ter Meulen (eds.), *The Representation of (In)definiteness*. Cambridge, MA: MIT Press, 98-129.
- Petinou K. & Terzi A. (2002). Clitic Misplacement among Normally Developing Children and Children with Specific Language Impairment and the Status of Infl Heads. *Language Acquisition* 10: 1-28.
- Philippaki-Warburton I. (1985). Word Order in Modern Greek. *Transactions of the Philological Society* 83: 113-143.
- Ralli A. (2006). Syntactic and Morphosyntactic Phenomena in Modern Greek Dialects: The State of the Art. *Journal of Greek Linguistics* 7: 121-159.
- Rizzi L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Rizzi L. (1997). The Fine Structure of the Left Periphery. In L. Haegeman (ed.), *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Kluwer, 281-337.
- Roussou A. & Tsimpili I.-M. (2006). On Greek VSO Again! *Journal of Linguistics* 42: 317-354.
- Simeonidis C. P. (2006). *Ιστορία της Κυπριακής Διαλέκτου* (History of the Cypriot Greek Dialect). Nicosia: Holy Monastery of Kykkos Centre of Research.
- Sinopoulou O. (2008). Multiple Questions and Apparent *Wh*-in situ: Evidence from Greek. *Proceedings of ConSOLE XV*: 223-246.
- Sinopoulou O. (2009). Απλές ερωτήσεις με ερωτηματική λέξη in situ: Η περίπτωση των Ελληνικών (Simple Questions with *Wh*-in situ: The Case of Greek). *Proceedings of the 8th International Conference of Greek Linguistics*: 1118-1132 (CD-ROM). [Also online at http://www.linguist-uoι.gr/cd_web/arxiki_en.html.]
- Sobin N. (2010). Echo Questions in the Minimalist Program. *Linguistic Inquiry* 41: 131-148.
- Starke, M. (2001). *Move Dissolves into Merge: A Theory of Locality*. Université de Genève doctoral dissertation, Geneva.
- Terzi A. (1999a). Clitic Combinations, Their Hosts and Their Ordering. *Natural Language and Linguistic Theory* 17: 85-121.
- Terzi A. (1999b). Cypriot Greek Clitics and Their Positioning Restrictions. In A. Alexiadou, G. Horrocks and M. Stavrou (eds.), *Studies in Greek Syntax*. Dordrecht: Kluwer, 227-240.
- Tsai W.-T. D. (2008). Left Periphery and *How-Why* Alternations. *Journal of East Asian Linguistics* 17: 83-115.
- Vlachos C. (2008). *Wh*-in-situ Inquiries in a *Wh*-Movement Language: The Case of Greek. Unpublished MS., University of Patras.
- Vlachos, C. (2010). *Wh*-in-situ: The Case of Greek. In V. Torrens, L. Escobar, A. Gavarró, and J. Gutiérrez (eds.), *Movement and Clitics*. Newcastle: Cambridge Scholars Publishing, 85-113.

Counterfactuality in the Tsakonian dialect: a contribution to the history of *ήθελα* and *ήμουν*

NIKOS LIOSIS

Institute of Modern Greek Studies
Aristotle University of Thessaloniki
nikosliosis@yahoo.gr

This article is concerned with some of the periphrastic conditional structures and wishes used by speakers of the Tsakonian dialect to express, as observed by Ferguson et alia (1986:3) describing these types of structures from a cross-linguistic perspective, conclusions based on inadequate information, imagined possible or alternative states of affairs, to understand the world when the relationships between things change.

According to Comrie (1986:88-9), on a continuum of hypotheticality, the lower the probability of realisation, the higher the degree of hypotheticality, and from this point of view, counterfactuals are located at one extreme of this continuum, having the highest possible degree of hypotheticality:

The continuum of hypotheticality



So, the unrealized and unrealizable conditionals and wishes, or, as Palmer (2001:207) puts it, those where the speaker shows some sort of negative attitude, are discussed here.

Traugott (1985· see also Lehmann 1974), in an attempt to define the universal markers of conditionals, identifies a very small number of types of non-conditional origin: a) modals of probability, doubt, wishing b) interrogatives c) copulas, usually of existential type d) topic markers and demonstratives e) temporals. Tsakonian makes use of the first and third options, as we shall see below. In this way the dialect differs from SMG (which uses the first option), not of course as regards the prototypical semantics of conditionality or the crosslinguistically established typology of conditionals, but rather as regards the lexical and morphological means chosen to express counterfactuality, and the morphosyntactic relationships established between these elements within the framework of grammaticalization theory.

For Lehmann (2002:29-30, 117-8), in periphrases which, as is usually the case with counterfactuals, are made up of two verbal elements, one of which is an auxiliary, in the first stages of grammaticalization the auxiliary governs, while, when its integrity has been eroded (for example with the loss of marking of certain verbal characteristics), it is the verb with lexical meaning which governs. This interpretation, seen from a comparative point of view, provides a useful typological schema for all varieties of Modern Greek based on two criteria proposed by Τζιτζιλής (forthcoming (a)):

a) The first criterion is modal past marking. The various varieties of Modern Greek may be divided into two groups: those with counterfactuals where the auxiliary is still marked for modal past, such as for example the dialects of Mykonos (Μάνεσης, 1997:348), e.g.

(1) /iθele na su dósi mila/ 'he would have given you apples'

of Chios (Pernot, 1946:289), e.g.

(2) /iθela yini foniko/ 'someone would have been murdered'

and of Avlonari (Φάβης, 1911:56), e.g.

(3) /iθela páis/ 'you would have gone',

and those with counterfactuals where the past is marked on the main verb, in other words where the counterfactual marker has undergone such a degree of phonetic reduction that it now coincides with the future marker; these varieties include, for example, SMG, the dialect of Corfu (Χυτήρης, 1992:233), e.g.

(4) /as ixa lefta ce tha m élyepes eména/ 'If I had had any money, you would have seen me'

and again Avlonari (Φάβης, 1911:56), e.g.

(5) /θela pijéname/ 'we would have gone'

etc.

The case of Avlonari is actually rather enlightening: given the fact that this dialect possesses the homophonous future marker /θela/, as seen in the future utterance (Φάβης, 1911:55)

(6) /θela γράψο/ 'I will write',

the reading of utterance (5) / θela pijéname/ as a counterfactual is secure, according to Τσολακίδης (2009:417), based only on the obligatory past tense of the main verb. In fact, as he observes, the crosslinguistic study by Bybbe, Perkins and Pagliuca (1994:515-6) has proved that in cases where the main verb is in the past, the counterfactual marker is the product of grammaticalization of the modal imperfect of the auxiliary (here /iθela/) rather than of the future marker (here /θelo/).

The difference between these two groups of varieties is not, according to Τζιτζιλής, simply a difference in the phonetic material of the marker, but is also semantic and grammatical. In the first group, the retention of past marking on the auxiliary allows the main verb to express clearly the distinction between perfective and imperfective aspect, such as in the two utterances from Demirdesi (Danguitsis, 1943:99-100)

(7) /θela γοράσο/ 'I would have bought',

(8) /θela γοράζο/ 'I would have been buying'.

On the contrary, in the group which includes SMG, the utterance θα αγοράζα is ambiguous: it can have perfective or imperfective meaning, past or future reference (see also Tomić, 2006:634-5) and habitual or non-habitual usage. In other words, as noted by Χόρροκς (2006:443), such constructions are neutral as regards both tense and aspect.

It is worth noting that in some cases, as for example that of utterances (3) and (5) from Avlonari, the material at our disposal from dictionaries, grammars and articles allows us to classify some dialects as belonging to both the groups defined above; this may be because we are dealing with constructions recorded at different chronological phases in

the development of the dialect, with differences between local subdialects, or it may be that the two constructions really did coexist as alternative possibilities for a certain period of time.

There are of course cases of intermediate / mixed dialects in which the past is marked on both the auxiliary and the main verb, either because they represent a transitional stage in the grammaticalization process, or as a result of influence from other language varieties. Examples of this type of “redundant” or transitional marking may be found in the dialect of Corfu (Χυτήρης, 1992:233), e.g.

(9) /iθela na sú dina mpa paraǰelia/ ‘I would have given you an order’,

and in Demirdesi (Danguitsis, 1943:100), e.g.

(10) /θela γόραza ένα άλογο/ ‘I would have bought a horse’.

b) The second criterion concerns the choice of auxiliary verb used in the periphrasis. The various varieties of Modern Greek select one of three auxiliaries in varying stages of grammaticalization, the most common being θέλω as in SMG, while others use έχω, e.g. Cypriot (Μενάρδος, 1925:45), e.g.

(11) /iǰen na γράpso/ ‘I would have written’,

the dialect of Kozani (Ντίνας, 2005:149), e.g.

(12) /xa na tun riksn pulés/ ‘they would have bitten him a lot’

and a small number, mainly from Asia Minor, use είμαι, such as the dialect of Axos (Μαυροχαλυβίδης - Κεσίσογλου, 1960:66), e.g.

(13) /na kréǰis ton/ ‘you would have searched’

and that of Silli (Κωστάκης, 1968:110), e.g.

(14) /itna su γράpsu/ ‘I would have written to you’

These introductory remarks will help us to more easily describe and interpret the equivalent Tsakonian constructions. According to the material at our disposal, which covers a time period extending from the mid-19th century to the present day, the two Tsakonian subdialects of the Peloponnese present a wide variety of different constructions, which, as I have already mentioned, include as counterfactual markers imperfect forms of θέλω and/or είμαι, and may be divided into the following categories:

1a) periphrastic auxiliary verb ήθελα + subjunctive

- /emaθa ráu/ (< */ema θέlu na ráu/ ≈ *ήμουν θέλων να οράσω) ‘I would have seen’
- /esaθa ráre/ ‘you would have seen’
- /eciθa rái/
- /emaiθa ráme/ etc.

A fundamental characteristic of the organisation of the Tsakonian verbal system is the periphrasticity of the present and imperfect tenses, which make use of the relevant tense of the stative auxiliary είμαι and the present participle, e.g. /emi γράfu/ (≈ *είμαι γράφων) ‘I write’ ~ /éma γράfu/ (≈ *ήμουν γράφων) ‘I was writing’. Within this framework the imperfect of the auxiliary verb θέλω is also constructed periphrastically, /ema θέ(l)u/, and is used in combination with both the perfective and imperfective subjunctive, cf. /emaθa orínu/ ‘I would have been seeing’. However, the use of a periphrastic verbal form in the construction of still more extended counterfactual structures increases their syntactic

complexity and constitutes a further source of pressure which encourages the operation of grammaticalization mechanisms, particularly those which lead to phonetic reduction of the material. These structures could be described as embedding periphrases (for the term see also Λιόσης, forthcoming), meaning that one periphrasis (here the imperfect) is incorporated as the first component, namely in a more grammatical position of a new periphrasis (here a counterfactual). Symmetrical with this and constructed in an equivalent manner is the future periphrasis of the type /emiθa ráu/ (= *είμαι θέλων να οράσω) ‘I will see’ (Λιόσης, forthcoming). In both cases the presence of the element /-θa/ could be considered the result of:

- a) a process of grammaticalization of the periphrastic θέλω which leaves the initial component, the inflected /éma/ (or /émi/ in the case of the future), unaffected, namely the deictic characteristics of tense, person and number (which are also “redundantly” marked on the lexical verb), but “erodes” the verb θέλω and the complementizer να, ultimately resulting in their coalescence (for the term see Lehmann, 2002:132): /θa/ (< /θa na/ < /θe na/ < /θeu na/), or
- b) the influence of the marker θα of SMG or neighbouring varieties which replaced the construction /θέλυ na/ following reanalysis and isolation of /ema/ as an autonomous element (for a more detailed discussion of this process, which also affects the future, see Λιόσης, forthcoming).

The structural model for the use of the inflected imperfect of θέλω together with the subjunctive may be traced back to the late mediaeval period. Markopoulos (2005:212) records a fairly large number of instances of the future-in-the-past from as early as the 15th century, such as the example (15) given below from Mahairas:

(15) *Είδα τον παπάν όπου εθέλα να κουρέψουν* ‘I saw the priest that they were about to consecrate’.

However, he emphasises the fact that until the 16th century, counterfactuals and conditionals occur exclusively with an infinitive complement, because evidently their grammatical context was particularly resistant to the syntactic development whereby the infinitive was replaced by complement clauses (see also Markopoulos, 2009:209-24).

Kostakis observes in addition the sporadic presence of other counterfactual markers deriving from the verb θέλω and να, which always appear in combination with the subjunctive (utterance 16 is from Southern Peloponnesian Tsakonian, 17-19 from the dialect of the Propontis):

a) /θala/ (1986 A':324):

(16) /θala záu ts ezú/ ‘I would have gone there, too’,

(17) /opsá na ta kanó, θala mi vrés/ ‘If he had come yesterday, he would have found me’.

b) /θela/ (1986 A':324, 327):

(18) /δέ θela éxoi ksíla na ksalíts^hoi/ ‘they wouldn’t have any wood left to burn’.

He observes that the particle also exists in Northern Tsakonian, but since he does not give examples it remains uncertain whether he is referring here to the future or to the counterfactual marker.

c) /θena/ (1986 A':324):

(19) /θe na spásoi ta kurbána/ ‘they were sacrificing’.

Kostakis considers that this particle too is also to be found in Northern Tsakonian, but since he does not provide us with examples it is again uncertain whether he is referring to the future or to the counterfactual marker.

d) /θewa/ (1956:125):

(20) /t abrésta θewa nápsoi ta tseria/ 'before that they were lighting candles'.

Utterances (19) and (20) are examples of the generic reading taken by counterfactuals with the subjunctive in past narratives (see also utterance (32), below). It is most likely that the form *θέουα* given by Kostakis in utterance (20) simply represents a phonetic variant of the type /θela/ showing the stage where intervocal /l/ was converted to a semi-vowel before its eventual deletion. If, however, we take the view that this form represents an earlier stage of grammaticalization (/θelu na/ > /θeu na/ > /θeu a/), we would be forced to accept the conclusion, improbable both from a theoretical point of view and with regard to the rules of this dialect, that at the initial stage of grammaticalization the masculine form of the participle (≈ *θέλων) is selected instead of the expected neuter (cf. 3rd person singular neuter participle in impersonal expressions such as /eni prépunda na záre ecu/ 'you (yourself) must go'). In any case, the first interpretation is also supported by the form /θea/ from example (30) below, which represents the final stage of the deletion of intervocal -l-.

For the following reasons the markers /θela/ (/θewa/, /θea/) ~ /θala/ ~ /θe na/ should in all probability be considered loans from the neighbouring Peloponnesian (see Pantelidis, forthcoming) or Bithynian dialects (see Τζιτζιλής, forthcoming (b)) and not as inherited Tsakonian:

- a) They replace the marker /ca/, which based on what we shall see represents the central element for the production of counterfactual structures in this dialect
- b) They lack the basic syntactic characteristic of periphrasticity, i.e. they are derived from a monolectic form of θέλω
- c) With regard to Propontis and Southern Tsakonian, with the exception of /θe(w)a/ they contravene the basic phonetic law of intervocal /l/-deletion, even if we accept that in the Propontis it is not applied as consistently as in Southern Tsakonian.

1b) periphrastic impersonal auxiliary verb ήθελε + subjunctive

-(e)ciθa ráu/ (< */eci θelu(nda) na rau/ ≈ *ήταν θέλ(ω/ο)ν να οράσω) 'I would have seen'

-(e)ciθa ráre/ 'You would have seen'

-(e)ciθa rái/

-(e)ciθa ráme/ etc.

We find the impersonal form of the auxiliary, /eciθa/, sporadically, most frequently in the northern Peloponnesian dialect, e.g.

(21) /eylitutse o papu o kakómere, p^hi ciθa i zemacisoi/ 'the poor old man was saved, or else they would have burned him'

(Λιόσης, 2007:452-3· for examples from the northern dialect see Κωστάκης, 1951:102). Such structures with the 3rd person form of the auxiliary which evidently constitute the starting point of the process of grammaticalization, as is generally considered to be the case also in SMG (see, for example, Χόρροκς, 2006:440-2), also correspond to similar structures found in the late mediaeval language and in other Modern Greek dialects which present a fossilized *ηθελε να* or *ηθελα < ηθελε να*, e.g.

(22) *ανέν και ηθελα λήπεις τότε* 'if you would be away at that time'

(Markopoulos, 2009:220· from the notary texts of Maras),

(23) /a den eruvárizes, iθele na peθáno/ ‘If you hadn’t come, I would have died’
(Mykonos· Μάνεσης, 1997:348),

(24) /as iθela me vuiθísis/ ‘I wish you had helped me’
(Eastern Crete· Πάγκαλος, 1955:329).

1c) periphrastic auxiliary verb *ήθελα* + marker-*είμαι* + subjunctive

-/emaθaca ráu/ (< */ema θelu na éci na rau/· ≈ *ήμουν θέλων να ήταν να οράσω) ‘I would have seen’
-/esaθaca ráre/ ‘you would have seen’
-/eciθaca rái/
-/emaiθaca ráme/ etc.

The syntactic length of the periphrasis increases still further with the presence next to the inflected θέλω of the marker /ca/, which is in its turn the product of the grammaticalization of the 3rd person form /eci/ ‘was’ and the marker /na/: /eci na/ > */ci na/ > */ci a/ > /ca/. The most characteristic point is the presence in the same construction of both auxiliaries, at different stages of grammaticalization. We will return to this.

1d) marker-*θέλω* + marker-*είμαι* + subjunctive

-/(e)θaca ráu/ (< */ema θelu na eci na rau/· ≈ *ήμουν θέλων να ήταν να οράσω) ‘I would have seen’
-/(e)θaca ráre/ ‘you would have seen’
-/(e)θaca rái/
-/(e)θaca ráme/ etc..

This option involves the phonetic reduction of the auxiliary /emi θέλυ/, or rather of the impersonal form /eci θα/: > /eḯ θα/ > /e θα/ (and > /θα/ as a result either of further phonetic reduction or of influence from SMG). The structure in question is reminiscent of the future periphrasis /e)θα ráu/ (< */eni θέλυ na ráu/) ‘I will see’, where the future marker has undergone the same degree of phonetic reduction; see Λιόσης, forthcoming. It is worth noting the existence of the rare form /θeca/ of the counterfactual marker along with /θaca/, e.g.

(25) /θeca ipoférume to molevo móre/ ‘We would have suffered on Malevos, dear’

(Λιόσης, 2007:808), which is reminiscent of equivalent dialect alternations *θελα* / *θαλα* or *θενα* / *θανα* (Peloponnesian· see Pantelidis, forthcoming). It is certainly the case that the presence of the element /θe-/ makes it more probable that the volitional itself (and not just its auxiliary) has passed through all the stages of grammaticalization, rather than being borrowed from SMG θα.

In addition, the use of the modal periphrastic imperfect in the formation of counterfactual structures, despite the fact that it invalidates the capacity of Tsakonian for distinction between perfective and imperfective aspect in the manner described above (cf. /θaca ráu/ ‘I would have seen’ ~ /θaca orínu/ ‘I would have been seeing’), creates two new syntactic options:

2a) marker-*θέλω* + imperfect

- /θα εμα ορύ/ (≈ *θα ήμουν ορών) 'I would have seen'
- /θα εσα ορύ/ 'you would have seen'
- /θα εσι ορύ/
- /θα εμαι ορύνδε/ etc.

This is a direct reflection of the usual SMG structure θα + imperfect, and is in all probability a case of borrowing of the SMG syntactic prototype.

2b) marker-θέλω + marker-είμαι + imperfect

- /θα ca εμα ορύ/ (≈ *θα ήταν να ήμουν ορών) 'I would have seen'
- /θα ca εσα ορύ/ 'you would have seen'
- /θα ca εσι ορύ/
- /θα ca εμαι ορύνδε/ κλπ.

This must be considered a hybrid form, since it appears to be a combination of the periphrases previously mentioned. The result is at first glance rather surprising, combining three modal markers, /θα/, /ca/ and the modal imperfect. It appears that these types of combinations are not exclusive to Tsakonian. In the dialect of Grevena we have the marker /xala/, which according to Τζιτζιλής and Μαργαρίτη-Ρόγκα (forthcoming; see also Τσολακίδης, 2009:418-9) resulted from the amalgamation of the auxiliaries /xana/ < /ίxa na/ and /θala/, as in the utterance

- (26) /an íksira xala páu ci iyú/ 'If I knew, I would have gone, too'

(Αναστασιάδης, 1998:17), while even closer to the Tsakonian pattern are mixed periphrases such as

- (27) ίψna (< ίψε na) ta páru ítu/ 'I would have taken them'

from the dialect of Silli (Κωστάκης, 1968:110), where the lexical verb is preceded by the 3rd person singular of έχω and followed by the 3rd person singular imperfect of είμαι.

3) marker-είμαι + subjunctive

- /ca ráu/ (= *ήταν να οράσω) 'I would have seen'
- /ca ráre/ 'you would have seen'
- /ca rái/
- /ca ráme/ κλπ.

The simplest but rarest form combines the indeclinable existential marker /ca/ with the subjunctive as in the following examples (both from Λιόσης, 2007:444):

- (28) /an éma kondá ta k^hára, ca fojistú/ 'If I had been near the fire, I would have been warmed'

- (29) /iɲɟaj ksérunde ots^hi ca móli o ts^hepéla/ 'they knew that Tsepela was about to come' (future-in-the-past)

This kind of periphrasis, however, is what links the two Peloponnesian subdialects with the Tsakonian subdialect of the Propontis: in the example

(30) /áma dé isa etu, thea peθán. o jéro dé ta borés na ftäs tiptaga/ 'If you hadn't been here (If it were not for you), I would have died; the old man could have done nothing'

(Κωστάκης 1986, Γ':423), the second counterfactual apodosis of the conditional is expressed with the use of the marker /ta/ (< /éta/ 'was' + /na/ 'to') and the subjunctive. Unfortunately, the fact that the main verb is in the 3rd person singular does not allow us to decide whether /ta/ remains inflected, but another example from Kostakis (1986, Α':192)

(31) /ná tai voleté, ma borés na paén/ 'If it had been convenient, I would have been able to go'

shows, although there is some doubt regarding the meaning, that this interpretation is indeed possible: /ma borés/ < /éma na borés/ (≈ *ήμουν να μπορέσω).

Moreover, the use of the structure /ta/ + subjunctive in narrative, where apparently conveys a generic meaning is very characteristic:

(32) /O κάθε spitonikots^hur ta pár éna petiné tse ta paén [...]. Tan tzefáa ta ni afis tsa péra tse ta pár ton petiné [...] tse tan áwa méra ta paénoi ston áje. [...] ta paén sto spiti s tse ta kasits na fái [...]./ 'Every house owner was taking a rooster and was going [...]. He was leaving the head there and was taking the rooster [...] and the next day they were going to church. [...] he was going home and was sitting down to eat.' (Κωστάκης, 1957:124 for equivalent generic uses of the structure θελα + subjunctive in Peloponnesian see Παντελίδης, forthcoming).

The presence of the marker of existential origin in the dialects of Asia Minor, among them the Tsakonian subdialect of the Propontis, could be interpreted as influence from Turkish, especially in dialects such as that of Axos (see utterance 13), where it follows the main verb. On the other hand, the choice of the same auxiliary for the formation of counterfactual periphrases in the dialect of Silli (see utterance 14) in all probability constitutes an isogloss linking this dialect with Tsakonian, lending support to the theory proposing a Tsakonian substrate in this region (for a more extensive discussion of the links between these dialects see Τζιτζιλής, forthcoming (c) and Τζιτζιλής, forthcoming (d)).

Whatever the case, the tendency to form future and consequentially counterfactual structures with verbs which mean 'be, become', which according to Bybee, Perkins & Pagliuca (1994:258-64) have their semantic starting point with meanings of obligation or predestination, is not found only in such exotic language varieties as Kui, Baluchi and Slave (258) to which the three authors refer. There is also a Balkan dimension to this phenomenon. In the western dialects of Slavic Macedonian, conditionals periphrases may be formed with the marker *bi*, which is derived from the Old Church Slavonic aorist *byti* 'I was', e.g.

(33) Ako bi da mu potrebvjet pari (= If + would + Subj. Mark.) 'If he happens to need money'

(Tomić, 2006:423, 444-5 and footnotes 64, 66). Similar structures are also found in SMG, e.g.

(34) *Είναι να πάω στο γιατρό / Ήταν να πάω στο γιατρό*

and can have readings which range from obligation to scheduled future, although they do not necessarily fulfil all the basic criteria to be considered periphrases (for these criteria see Aerts, 1965:3; Haspelmath, 2000:654-5).

In Lehmann's terminology (2002:120-1), the two counterfactual markers used in Peloponnesian Tsakonian show the highest degree of paradigmatic integration, given that

they even combine with perfect tenses. These structures are of three types (Λιόσης, 2007:443):

4a) per. aux. verb ήθελα + present perfect subj.	4b) marker-θα + marker-ca + present perfect subj.	5) marker-θα + past perfect
-/εμαθα exu orate/ (≈ *ήθελα να έχω ορατό) 'I would have seen'	-/θα ca exu orate/ (≈ *θα ήταν να έχω ορατό) 'I would have seen'	-/θα εμα exu orate/ (≈ *θα ήμουν έχων ορατό) 'I would have seen'
-/εσαθα εçere orate/ 'you would have seen'	-/θα ca εçere orate/ 'you would have seen'	-/θα εσα exu orate/ 'you would have seen'
etc.	etc.	etc.

In utterances of type 4a, as well as utterances such as the following from other dialects, e.g.

(35) /θelana ts éxum mazuménis tsi les/ 'we would have gathered the olives'
(Σαμοθράκη· Τσολάκη, 2009:425),

(36) /iθena tó xo vγáli/ 'I would have removed it'
(Κίμωλος· Βογιατζίδης 1925:157),

(37) /an ienná çis féri tok K, ienná rti c o A/ 'If you had brought K., A. would have come, too'

(Κάρπαθος· Μηνάς, 1970:109),

the ambiguity between the past and future readings is resolved in favour of the former with the combination of the modal past of the auxiliary and the perfect aspect of the lexical verb. Conversely, in 5, which, like the structure with the main verb in the imperfect (see 2a and 2b), must be considered a loan from SMG, the grammaticalization of the auxiliary to the point where its past tense origin is obscured, creates a need for double marking of the main verb as regards time reference: 'once for the past, once for unreality', as Palmer characteristically observes regarding equivalent structures in English such as the protasis in the utterance 'If John had come, Bill would have left' (2001:208; note the equivalence between the material used to form the apodosis in the English utterance, 4a, and the utterances from Samothrace, Kimolos and Karpathos); the only "doubly" past tense is of course the past perfect, described by Tomić (2006:633) in combination with θα as "future past-perfect-in-the-past". In Tsakonian, however, we also find the option 4b: the marker /ca/ is retained only when there is no past marking on any of the other components of the periphrasis, and the same applies in the case of the utterances /θaca ráu/ ~ /θaca orínu/ (see above). In other words, it is the marker /ca/ which prevents synonymy with the futures /θα éxu oraté/ 'I will have seen', /θα ráu/ 'I will see' and /θα orinu/ 'I will be seeing' respectively.

If we attempt a relative chronology of the two markers, we may conclude that /ca/ is older based on the following observations:

a) Unlike θέλω, which can be inflected, είμαι always appears completely grammaticalized, which allows us to suppose that it has been in use as a counterfactual marker for a longer period of time.

b) It has a «harmonic» presence (for the Harmony principle, see Bybee, Perkins & Pagliuca, 1994:214-225) in the protasis of conditional structures, which, «being just as modal» as the apodosis, as noted by Horrocks (2006:439) «eventually make use of the same forms», e.g.

(38) /naca mólere, θaca nd oráu/ 'If you had come I would have seen you'

(Κωστάκης, 1986 B':291), while it also appears frequently in negation environments:

(39) /óca (< /ú éci na/ 'not was to') bret^hú/ 'I would not have been wet'

(Λιόσης, 2007:445). Its presence therefore in such syntactic environments, which are either conservative, such as negation (see Givon, 1979a [in: Bybee, Perkins & Pagliuca, 1994:237] and Givon, 1994 as well, for the conservative nature of negation), or non-assertive, such as subordinate clauses, may be taken as proof that it is old; note the equivalent "old" structure, which is in fact also formed using the existential verb, in the protasis of the following conditional from western Crete:

(40) /ná tone ná xo, iθela su dóso/ 'If I had had (it) I would have given (it) to you'

(Πάγκαλος, 1955:330); cf. Τσολακίδης (2009:423) and Τσολακίδης (forthcoming) for the relative chronology of the auxiliary έχω.

c) Its syntactic position is always closer to the lexical verb than that of θέλω, namely it constitutes the nucleus of the tripartite periphrasis.

Finally, mention must be made of another periphrasis which can be compared to structure 1d) above, and which is more common in Northern than in Southern Tsakonian:

6) marker-θέλω + marker-είμαι + "bare" subjunctive

-/θaci ráu/ (< */ema θelu na eci rau/ ≈ *ήμουν θέλων να ήταν οράσω) 'I would have seen'

-/θaci ráre/ 'you would have seen'

-/θaci rái/

-/θaci ráme/ etc.

Kostakis, in his grammar of the Northern subdialect (1951:102) includes the declensional paradigm:

(41) /θa ci fténu/, /θa ci fténe/ etc. 'I would have baked, you would... etc.'

and utterances such as:

(42) /θa ci s plerúi/ 'he would have paid for them' etc.

In his dictionary (1986, A':286) he gives a further example, this time from Southern Tsakonian:

(43) /eréste a elía purtése aiδέ θa ci zái t^ho gatáva/ 'The olive tree happened to be in his way, otherwise he would have gone down (fallen off the cliff)'

Here we have a combination of impersonal /eci/ with a verb in the subjunctive without the presence of the complementizer /na/, i.e. the clausal complement is replaced by the "bare" subjunctive. Such constructions are not unknown in the history of the Greek language. Markopoulos (2009) refers to the existence of future and counterfactual structures with θέλω + subjunctive without να in the late mediaeval period (166-7 and 220), stating emphatically that since these types of structures also occur with έχω (71-2) and μέλλω (128-9), the other two auxiliaries that historically have given future and counterfactual structures, it is impossible that the model they represent could have developed from periphrases with θέλω + infinitive, as they were until now believed to have done (see, for example, Joseph & Pappas, 2002: Χόρροκς, 2006:440-1). The examples

from Tsakonian are particularly helpful with regard to this matter: they show that structures with the bare subjunctive are also found using the fourth auxiliary, είμαι, thus completing the picture and supporting Markopoulos's argument. If we accept that this syntactic phenomenon is very old, dating back as far as Ancient Greek (Markopoulos, 2009:38-9), we have yet another argument indicating that the use of είμαι predates that of θέλω in the formation of modal periphrases in Tsakonian. More generally, it may be concluded that the study of the Modern Greek dialects can be extremely useful in determining the correct chronological, geographical and theoretical basis for the discussion of such issues. For example, the presence of the same type of perfect structures in the dialect of Corfu, e.g.

(44) /έxo fáo, έçis fáis, έçi fái, έxume fáme/ etc. 'I have eaten, you...etc.'

(Κρίκη & Λιόσης, forthcoming) shows that the issue at hand in fact affects the whole system of moods, tenses and aspects in Greek.

Conclusions

The coexistence of the two markers in counterfactual periphrases should certainly not be considered a case of unmotivated accumulation. The most probable interpretation is that extensions of the use of /ca/ gradually obscured its function as a counterfactual marker. This function was reinforced by the addition of the imperfect of θέλω, which was grammaticalized in its turn. This cycle of feedback between the introduction of past tense elements and their subsequent grammaticalization was completed with the introduction of a third past marker, the modal imperfect of the main verb. That the successive modal markers were introduced in this particular order (rather than for example an earlier use of the imperfect) is confirmed by the complete absence of structures combining /ca/ by itself with the imperfect, e.g. */ca ema oru/ (= *ήταν να ήμουν ορώv). That counterfactual markers are often subject to this kind of reinforcement is nothing new in the bibliography: Dahl (1997:109) observes that the need for emphasis plays an important role: "Markers of hypotheticality might originate with locutions that are used to underscore the falsehood of an assumption and are later subject to extensions in their use and simultaneous weakening of their force. This in its turn may lead to the rise of new markers, and another round in the cycle." It is simply that in Tsakonian, the appearance of each new marker was not necessarily accompanied by the loss of its predecessor (cf. the English future, which today may be formed with *will*, *shall* or *be going to* (Bybee, Perkins, Pagliuca, 1994:21).

The reasonable hypothesis that sometime in the near past the choice between the two markers lead to distinctions of semantics, style or pragmatics (e.g. distinctions on the continuum of conditionality or of time reference in relation to the moment of utterance) remains unconfirmed, given that today the two elements, even when they are not used in the same periphrasis, are equivalent in meaning and distribution, and may even be found in the same utterance, e.g.

(45) /θa ca móli tatsip^héri to kabzi, θa émai aúde re jórgo/ 'the child would have come the day before yesterday, we would have talked with him, George'

(Λιόσης, 2007:808-9).

Τζιτζιλής (forthcoming (a)), however, observes that dialects which preserve synchronically different degrees of grammaticalization of the auxiliary (or of different auxiliaries) are able, by changing or specializing their meaning, to express detailed distinctions on the continuum of hypotheticality for example *potentialis* between *realis* and *irrealis*. Whether or not dialects possess the capacity to do this could be used as a third criterion for grouping them.

The above analysis reveals that, according to the model discussed at the beginning of the paper, Tsakonian presents a mixed typology as regards the distribution of modal past

marking and as regards the choice of auxiliary. Two counterfactual markers showing a greater or lesser degree of grammaticalization, functionally interchangeable, which may coexist in the same periphrasis and, being hypercharacterized in comparison with SMG, may combine with three aspects (perfective, imperfective, perfect) and two past tenses (imperfect, past perfect), certainly could not be called a prototypical case, and this demonstrates once again the unique character of this dialect.

References

- Aerts J. (1965). *Periphrastica*, Hakkert, Amsterdam.
- Αναστασιάδης Β. (1998). Μερικοί γραμματικο-συντακτικοί ιδιοματισμοί στο γρεβενιώτικο νεοελληνικό ιδίωμα, *Νεοελληνική Διαλεκτολογία* 2: 11-21.
- Βογιατζίδης Ι. 1925. Περί της Γλώσσης των Κιμωλίων, *Αθηνά* 36: 109-60.
- Bybee J., Perkins R. & Pagliuca W. (1994). *The evolution of grammar: tense, aspect and modality in the languages of the world*, U.C.P., Chicago.
- Comrie B. (1986). Conditionals: a Typology. In E. Traugott, A. Meulen, J. Reilly and Ch. Ferguson (eds) *On Conditionals*, Cambridge University Press, Cambridge, 77-99.
- Dahl Ö. (1997). Past time reference and counterfactuality. In A. Athanasiadou & R. Dirven (eds) *On Conditionals again*, John Benjamins Publishing Company, Amsterdam / Philadelphia, 97-114.
- Danguitsis C. (1943). *Étude descriptive du dialecte de Démirdési*, Maissonneuve, Paris.
- Ferguson C., Reilly J. Meulen A. & Traugott E. (1986). *Overview*. In Traugott E., A. Meulen, J. Reilly and C. Ferguson (eds) *On Conditionals*, Cambridge University Press, Cambridge, 1-20.
- Givon T. (1994). Irrealis and the subjunctive, *Studies in Language* 18: 265-337.
- Haspelmath M. (2000). Periphrasis. In G. Booij, V. Lehmann and J. Mugdan (eds.), *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*, DeGruyter, Berlin.
- Joseph B. & Pappas P. (2002). On some recent views concerning the development of the Greek future system, *BMGS* 26: 247-73.
- Κρίκη Ε. & Λίσης Ν. (forthcoming). Τα επτανησιακά ιδιώματα. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Κωστάκης Θ. (1951). *Σύντομη γραμματική της τσακωνικής διαλέκτου. Α. Το ιδίωμα της βόρειας Τσακωνιάς. Β. Το τσακόνικο ιδίωμα της Προποντίδας*, Institut Français d'Athènes, Αθήνα.
- Κωστάκης Θ. (1957). Τσακόνικα παραμύθια, *Λαογραφία* 17: 93-178.
- Κωστάκης Θ. (1968). *Το γλωσσικό ιδίωμα της Σίλλης*, Κέντρο Μικρασιατικών Σπουδών, Αθήνα.
- Κωστάκης Θ. (1986). *Λεξικό της τσακωνικής διαλέκτου*, ν. Α', Β', Γ', Αθήνα, Ακαδημία Αθηνών.
- Lehmann C. (1974). A universal about conditional sentences, in: Romportl, M. et al. (eds.), *Linguistica generalia I. Studies in linguistic typology*. Prague: Charles University (Acta Universitatis Carolinae, Philologica 5): 231-241.
- Lehmann C. (2002). *Thoughts on grammaticalization*, Universität Erfurt, Erfurt.
- Λίσης Ν. (2007). *Γλωσσικές επαφές στη νοτιοανατολική Πελοπόννησο*, Διδακτορική Διατριβή, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.
- Λίσης Ν. (forthcoming). Η γραμματικοποίηση στις περιφραστικές δομές μέλλοντα της τσακωνικής: ιστορικές και κοινωνιογλωσσικές παρατηρήσεις, *ΜΕΓ* 30, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Μάνεσης Σ. (1997). *Ιστορικό λεξικό του μυκονιάτικου ιδιώματος*, Δήμος Μυκονίων, Μύκονος.
- Markopoulos T. (2005). *The Category of Future in Greek*, PhD Dissertation, University of Cambridge, Cambridge.
- Markopoulos T. (2009). *The Future in Greek. From Ancient to Medieval*, Oxford University Press, Oxford.
- Μαυροχαλυβίδης Γ. & Κεσίσογλου Ι. (1960), *Το γλωσσικό ιδίωμα της Αξού*, Collection de l'Institut Français d'Athènes, Αθήνα.
- Μενάρδος Σ. (1925). Κυπριακή Γραμματική Γ, *Αθηνά* 37: 35-79.
- Μηνάς Κ. (1970). *Τα ιδιώματα της Καρπάθου*, Διδακτορική Διατριβή, Αθήνα.
- Ντίνιας Κ. (2005). *Το γλωσσικό ιδίωμα της Κοζάνης*, ν. Α', Κοζάνη.
- Palmer F. (1986), *Mood and Modality*, Cambridge: Cambridge University Press, 2001.
- Πάγκαλος Γ. (1955). *Περί του γλωσσικού ιδιώματος της Κρήτης*, 1st v., Αθήνα.
- Παντελίδης Ν. (forthcoming). Πελοποννησιακή. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Pernot H. (1946). *Morphologie des parlars de Chio*, Les Belles Lettres, Paris.

*Counterfactuality in the Tsakonian dialect:
a contribution to the history of 'ήθελα' and 'ήμουν'*

- Τζιτζιλής Χ. (forthcoming a). Εισαγωγή. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τζιτζιλής Χ. (forthcoming b). Τα θρακοβιθυνιακά ιδιώματα. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τζιτζιλής Χ. (forthcoming c). Τα λοιπά μικρασιατικά ιδιώματα. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τζιτζιλής Χ. (forthcoming d). Η τσακωνική διάλεκτος. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τζιτζιλής Χ. & Μαργαρίτη-Ρόγκα Μ. (forthcoming). Τα βόρεια ιδιώματα. In Χ. Τζιτζιλής (ed.), *Οι νεοελληνικές διάλεκτοι*, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τομιό Ο. (2006). *Balkan Sprachbund. Morphosyntactic features*, Springer, Dordrecht, 2006.
- Τσολάκη Μ. (2009). *Μορφολογία του Ιδιώματος της Σαμοθράκης: Ονοματικό και ρηματικό κλιτικό σύστημα*, Διδακτορική Διατριβή. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.
- Τσολακίδης Σ. (2009). Αντιγεγονοτικές περιφράσεις με το ήθελα να στις ΝΕ διαλέκτους, *ΜΕΓ* 29: 413-427, Ινστιτούτο Νεοελληνικών Σπουδών (Ίδρυμα Μανόλη Τριανταφυλλίδη), Θεσσαλονίκη.
- Τσολακίδης Σ. (forthcoming). Γραμματικοποιημένες δομές δυνητικής τροπικότητας στις νεοελληνικές διαλέκτους. *3ο Διεθνές Συνέδριο Νεοελληνικών Διαλέκτων και Γλωσσολογικής Θεωρίας*, Λευκωσία, 14-16/6-2007.
- Traugott E. (1985). On conditionals. In J. Haiman (ed.), *Iconicity in Syntax*, Benjamins, Amsterdam, 289-307.
- Φάβης Β. (1911). *Γλωσσικά επισκέψεις αναφερόμενοι εις το γλωσσικόν ιδίωμα Αυλωναρίου και Κονίστρων*, Λεωνής, Αθήνα.
- Χόρροκς Τ. (2006). *Ελληνικά. Ιστορία της γλώσσας και των ομιλητών της*, Introduction-translation in Greek: Μ. Σταύρου and Μ. Τζεβελέκου, Εστία, Αθήνα.
- Χυτήρης Γ. (1987). *Κερκυραϊκό γλωσσάρι*, ΕΚΣ, Κέρκυρα, 1992.

Towards the acoustic analysis of lateral consonants in Modern Greek dialects: a preliminary study

ANASTASSIA LOUKINA
Oxford University Phonetics Laboratory
anastassia.loukina@stx.ox.ac.uk

1. Introduction

Published phonetic descriptions of Modern Greek dialects report substantial variation in lateral consonants. Most if not all Modern Greek varieties appear to have several non-contrastive voiced lateral approximants whose distribution is determined by the following vowel. In many varieties palatal [ɲ] or palatalised dental [ɲʲ] occur before front vowels. In Northern dialects palatal consonant may also occur in word-final position or before a consonant as a result of deletion of etymological unstressed /i/. According to Arvaniti (1999b) in Cyprus speakers may use [j] instead of [ɲ]. Similar pronunciation is attested in some Chios villages (Kontosopoulos 2001). In several varieties including Standard Modern Greek palatal consonant may also occur before back vowels: Ath. [maɲˈla] 'hair'. These cases are usually analysed in studies on Modern Greek phonology as an underlying sequence /li/. The only exception to this trend is Tsakonian where according to Kontosopoulos (Kontosopoulos 2001) /l/ is velarised before [i].

In most part of Epirus, Macedonia and other Northern Greek dialects as well as in Western Crete lateral consonants before back vowels are velarised. Joseph and Tserdanelis (2003) note that this pronunciation serves as a regional identifier for northern speakers. In some varieties including Naxos and Crete velarised [ɲˠ] had been vocalized into [w].

Furthermore, in Cypriot Greek and other South-Eastern dialects all sonorants including lateral consonants have phonetically long counterparts usually called 'geminate'. Geminate sonorants occur word-medially and word-initially and may be lexical or post-lexical. In some regions of Crete and some Dodecanese dialects words which elsewhere have [l:] are pronounced with [lt] or [ld], while words that elsewhere are pronounced with [l] can be pronounced with an approximant [ɲ] (Kontosopoulos 2001; Joseph and Tserdanelis 2003).

This brief overview shows the width of the reported variation. Yet, most of these descriptions are based on impressionistic observations, and duration and quality of lateral consonants are among the features that constitute a particular challenge for an impressionistic auditory analysis. Therefore an instrumental analysis may substantially improve their description. It is also known from acoustic studies on lateral consonants in other languages that lateral consonants are generally subject to contextual variation, co-articulation and individual variation (Ladefoged and Maddieson 1996). This raises questions about the consistency and scope of the reported phenomena, which can only be answered by a quantitative instrumental study.

Instrumental data on lateral consonants in Modern Greek dialects are only available for Cypriot Greek geminates and laterals in Patras dialects. These studies revealed new aspects of variation in lateral consonants. For example, Papazachariou (2003) showed that in the dialect of Patras regional (more palatalised) pronunciation of /l/ was in free variation with the standard pronunciation. Eftychiou (2008) showed that geminate laterals in Cypriot Greek differed from corresponding singletons not only in duration (see Arvaniti 1999a, Tserdanelis & Arvaniti 1999, Arvaniti & Tserdanelis 2000, Arvaniti 2001), but also by a consistent differences in quality. She also found that geminate laterals involved a greater amount of linguo-palatal contact, especially word-initially.

In this paper I will look at quality and duration of [l] and [l:] in the three varieties of Greek: Thessalian, Cypriot and Athenian Greek. Thessalian Greek is an example of the Northern dialects and like other Northern Greek dialects, is reported to have ‘dark’ or ‘velarised’ [l^v] before back vowels (Kontosopoulos 2001). Cypriot Greek is an example of the South-Eastern dialects, which distinguish between geminate and singleton laterals. The third variety included in this study is Athenian Greek, which was chosen in order to provide some benchmark data that would be as close as possible to a natural colloquial form of Standard Modern Greek.

The analysis has several goals: to provide a more precise phonetic description of regional features and possibly reveal further differences between the dialects; to separate the truly regional features from processes which occur in colloquial Greek elsewhere; and to compare the patterns of variation between the varieties. I will examine whether spontaneous speech in Cypriot Greek supports the findings obtained on laboratory speech and compare the quality and duration of geminate and singleton laterals in Cypriot Greek to laterals in Athenian and Thessalian Greek. I will also investigate whether the impression of ‘dark’ [l^v] reflects the difference in acoustic properties of Thessalian /l/ from /l/ before back vowels in the other two dialects. Finally, I will look at what other factors may affect the duration and quality of lateral consonants in all three varieties.

2. Data and measurements.

The study is based on the same corpus as previously described in (Loukina in press). It consists of spontaneous monologues recorded from 21 speakers in Athens, Thessaly (Karditsa) and Cyprus (Nicosia). All speakers belonged to the same age group (75-93 years old), had primary education and were involved in traditional occupations. All speakers from Cyprus and Thessaly were natives of the area. Speakers recorded in Athens lived there at least since 1950s and were not perceived as regional speakers by speakers of Standard Modern Greek.

The analysis is based on the comparison of /l/ (/l:/) in 225 tokens of three highly frequent words shown in

Table . In all cases the lateral consonant is in pre-stress position word-medially, surrounded by back vowels.

Table 1. *List of tokens with Athenian, Thessalian and Cypriot pronunciation, Standard Greek spelling and English translation.*

Token	Athenian	Thessalian	Cypriot	Greek spelling	English
<i>kala</i>	[ka ^l a]	[ka ^l a]	[ka ^l a]	καλά	well, good
<i>polla</i>	[pɔ ^l a]	[pu ^l a]	[pɔ ^l :a]	πολλά	many
<i>poli</i>	[pɔ ^l i]	[pu ^l i]		πολύ	a lot

The geminate lateral in Cypriot [pɔ^l:a] is lexical and is reflected in the spelling. In Cypriot Greek [pɔ^l:a] is used for Standard Modern Greek [pɔ^li]. Although some of the speakers occasionally used the standard form [pɔ^li], the number of occurrences was low and therefore the Cypriot data for this word was not included into the analysis. In Thessalian Greek etymological unstressed /o/ has the same distribution of frequencies of F1 as the stressed /u/ and significantly higher than the F1 frequency of unstressed /o/ in Athenian or Cypriot Greek (Loukina in press).

All tokens of these words were saved as separate sound files and analysed using Wavesurfer¹ speech processing software. The tokens were manually segmented into phones. For the purpose of this study following Peterson and Lehiste (1960) /l/ was

¹ <http://www.speech.kth.se/wavesurfer/>

identified based on the changes in formant frequencies and amplitude. The durations were automatically extracted from the labels. Formants were tracked using the formant-tracking function of the software and manually checked against the spectrogram for accuracy. The frequencies of formants were automatically extracted at the interval of 10 ms and the value closest to the middle of the segment was used for further analysis.

Several normalization procedures were applied, including average values discussed by Adank (2004) and z-scores as proposed by Lobanov (1971). None of them allowed differences between speakers to be removed while preserving contrast between different sounds. Therefore it was decided in the first instance to use the raw frequencies in Hz, dealing with variation due to speaker and dialect via the statistical tests employed rather than normalization.

3. Duration of lateral consonants in Modern Greek dialects.

Table and Figure show durations of lateral consonants in the three varieties of Greek. A Mann-Whitney U test showed that only in Cypriot Greek was there a consistent significant difference in the duration of /l/ between *polla* and *kala* (120 ms vs. 63 ms, $p < 0.001$). This agrees with previous accounts of Cypriot geminates, which showed that geminate laterals in Cypriot are usually longer than the corresponding singletons.

Table 2: Mean duration of /l/ (ms) in *polla* and *kala* in Thessalian, Cypriot and Athenian Greek. The numbers in italics indicate standard deviation.

	<i>polla</i>	<i>kala</i>	<i>poli</i>
Athenian Greek	86.7	74.9	74.1
	<i>20.5</i>	<i>12.6</i>	<i>0.23</i>
Thessalian Greek	66	64.3	63.8
	<i>21.1</i>	<i>13.2</i>	<i>0.24</i>
Cypriot Greek	119	63.3	-
	<i>33.9</i>	<i>19.2</i>	-

Contrary to the results reported by Arvaniti (1999a) based on laboratory speech, in spontaneous speech there was certain overlap in duration between singleton and geminate consonants, but they had very distinct distributions of duration.

Comparison between the varieties showed that there also was significant difference in duration of the lateral consonant in *polla* between Cypriot Greek and the other two varieties (119 ms in Cypriot Greek vs. 66 ms in Thessalian Greek and 87 ms in Athenian Greek, Mann Whitney U tests, $p < 0.001$), but no significant differences between the durations of /l/ in *kala*. Furthermore, even though Athenian and Thessalian Greek do not distinguish between geminates and singletons, they did not show greater variation in duration than Cypriot singletons.

There also was no difference in duration between /l/ followed by back vowels and /l/ followed by front vowel in Athenian and Thessalian Greek.

4. Quality of lateral consonants in Modern Greek dialects.

The frequencies of F1 and F2 of /l/ in the three varieties are shown in *Table* .

Table 3: Mean F1 and F2 frequencies of /l/ (Hz) in *polla*, *kala* and *poli* in Athenian, Thessalian and Cypriot Greek. The numbers in *italics* indicate standard deviation.

	<i>polla</i>		<i>kala</i>		<i>poli</i>	
	F1	F2	F1	F2	F1	F2
Athenian Greek	532	1490	571	1441	335	1690
	<i>159</i>	<i>196</i>	<i>159</i>	<i>235</i>	<i>72</i>	<i>246</i>
Thessalian Greek	465	1324	561	1356	279	1839
	<i>80</i>	<i>232</i>	<i>100</i>	<i>253</i>	<i>42</i>	<i>301</i>
Cypriot Greek	355	1448	480	1474		
	<i>122</i>	<i>153</i>	<i>130</i>	<i>123</i>		

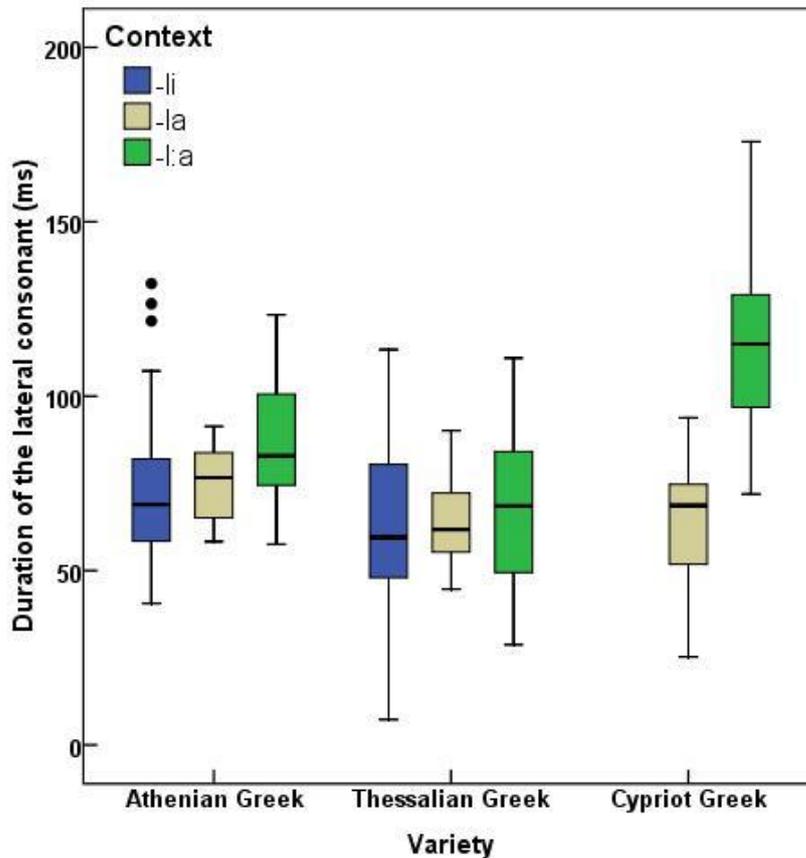


Figure 1: Distribution of durations of /l/ in *kala* and *polla* in Athenian, Cypriot and Thessalian Greek. The boxes show the data between the 25th and 75th percentile, the band near the middle of each box indicate median value. The whiskers indicate the lowest and highest datum within 1.5 of the interquartile range. Black dots indicate the outliers.

4.1. Singletons and geminates.

In Cypriot and in Thessalian Greek there was a significant difference in F1 of /l/ between *polla* and *kala* (355 Hz vs. 480 Hz in Cypriot Greek, 465 vs. 561 Hz in Thessalian Greek, Mann-Whitney U tests in both cases, $p < 0.001$). In both these varieties the /l/ in *polla* has a lower F1 than in *kala* (see *Figure*). The results for Cypriot Greek correspond to the results obtained on laboratory speech by Eftychiou (2008). She has shown that in

Cypriot Greek word-medial geminate /l:/ has a lower F1 and a tendency towards lower F2 than the corresponding singleton. The results of this study confirm that the difference in F1 between singletons and geminates is also present in spontaneous speech; however, in this data sample there was no difference in F2.

In Thessalian Greek /l/ in *polla* had lower F1 than in *kala*. There is no other evidence for contrast between geminate and singleton laterals in Thessalian Greek and there was no difference in duration between lateral consonants in these two words. Therefore it is unlikely that this difference can be explained by the lexical contrast as in case of Cypriot Greek. The more likely explanation is the influence of the preceding vowel. Unlike Athenian Greek, where /l/ is preceded by [ɔ], in Thessalian Greek /l/ is preceded by [u]. This could have resulted in greater difference in F1 between /l/ in *polla* and in *kala* in Thessalian Greek than in Athenian Greek. I will discuss the effect of adjacent vowels later in this paper.

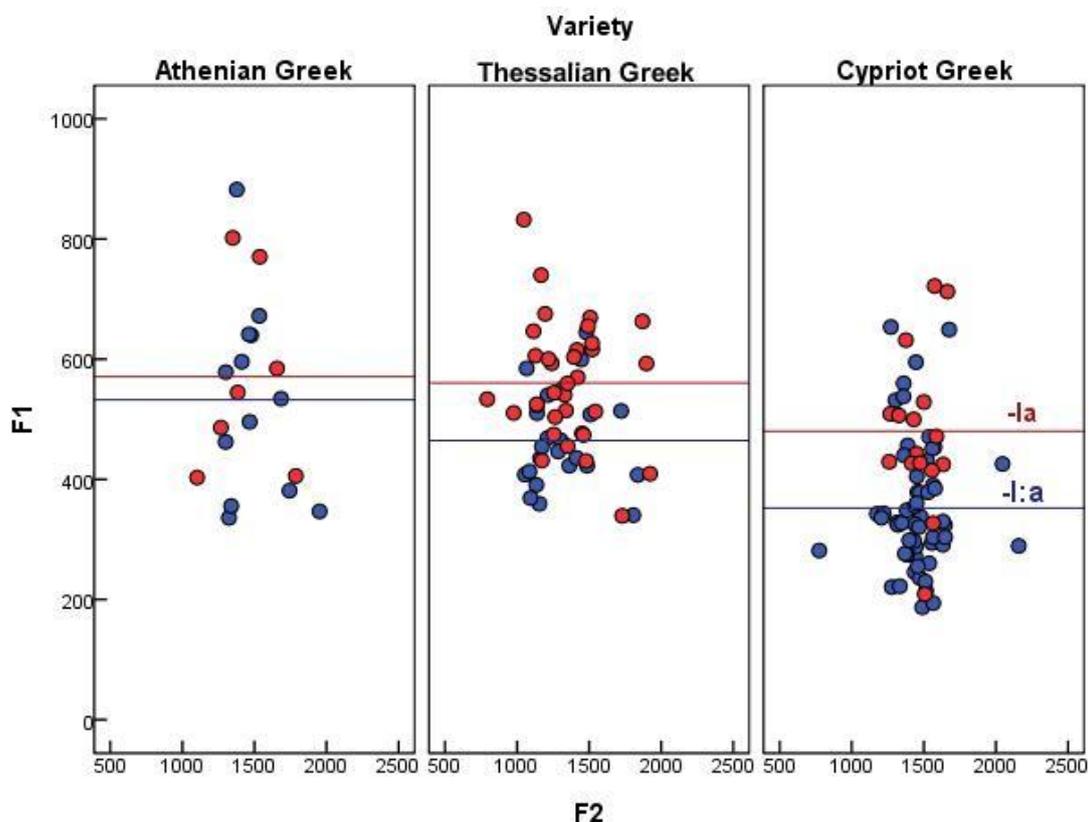


Figure 2: The frequencies of F1 and F2 (Hz) of /l/ in *kala* (red) and *polla* (blue) in Athenian, Thessalian and Cypriot Greek. Horizontal lines indicate mean values for each variety.

Comparison between the varieties showed that lateral consonants in Cypriot Greek had lower F1 than in Athenian and Thessalian Greek. The geminate /l:/ in Cypriot Greek *polla* had a lower F1 than the singleton /l/ in the same word in Athenian and Thessalian Greek (255 Hz in Cypriot Greek vs. 465 Hz in Thessalian Greek and 532 Hz in Athenian Greek, Mann-Whitney U tests, $p < 0.001$). For singletons there was a difference in F1 between Thessalian and Cypriot Greek, with /l/ in Cypriot having a lower F1 (561 Hz vs. 580 Hz, Mann-Whitney U test, $p < 0.05$).

4.2. Quality of lateral consonants before back and front vowels

Comparison of quality of /l/ before front and back vowels in Athenian and Thessalian Greek² showed significant differences in frequencies of both formants depending on the following vowel.

In both varieties the lateral consonant had lower F1 and greater F2 before /i/ than before /a/ (Mann-Whitney U tests in all cases, $p < 0.001$. See *Table* and *Figure*). In *poli* /l/ in Thessalian Greek had higher F2 (1839 Hz vs. 1690 Hz, Mann-Whitney U test, $p < 0.05$) and lower F1 than in Athenian Greek (279 Hz vs. 335 Hz, Mann-Whitney U test, $p < 0.01$).

Comparison between the varieties showed that in Thessalian Greek /l/ had a lower F2 than in Athenian or in Cypriot Greek in *polla* (1324 Hz vs. 1490 Hz in Athenian Greek, Mann-Whitney U test $p < 0.05$, and vs. 1448 Hz in Cypriot Greek, $p < 0.001$). The /l/ in *kala* in Thessalian Greek had a lower F2 than in Cypriot Greek (1356 Hz vs. 1474 Hz, Mann-Whitney U test, $p < 0.05$), but there was no significant difference from Athenian Greek (possibly due to the low number of Athenian tokens).

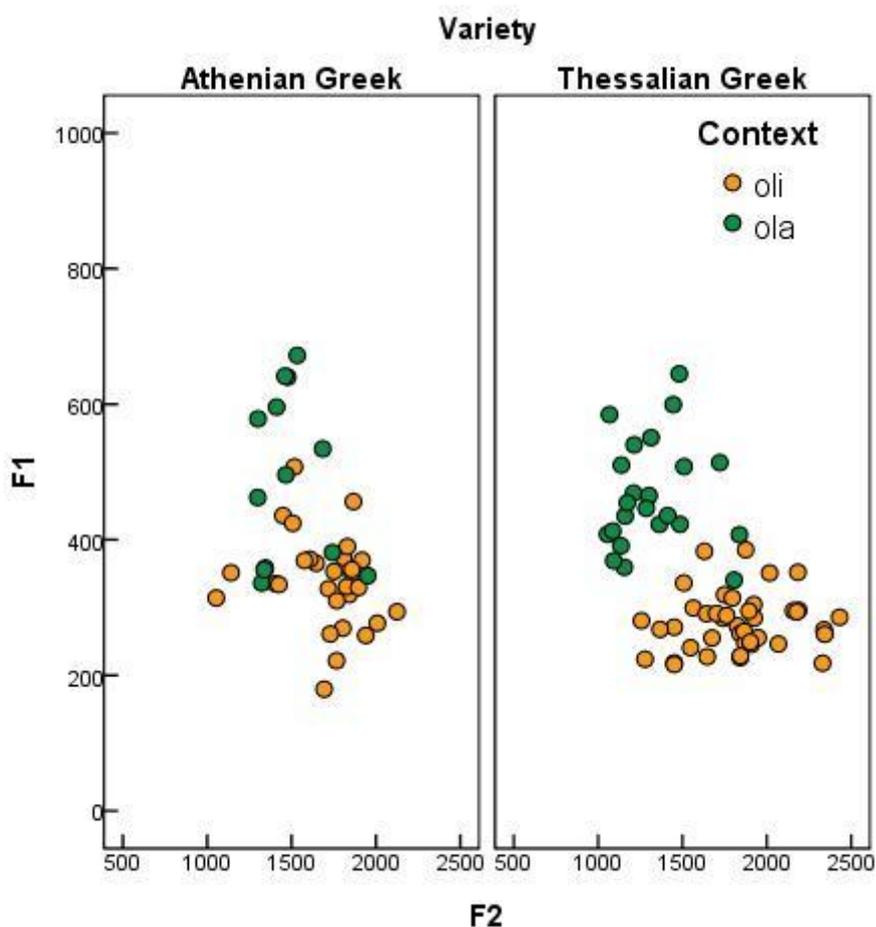


Figure 3: The frequencies of F1 and F2 (Hz) of /l/ in *poli* (yellow) and *polla* (green) in Athenian and Thessalian Greek.

Lower F2 in lateral consonant may indicate velarisation (Ladefoged and Maddieson 1996). Therefore results of this study agree with the impressionistic observation that Thessalian Greek /l/ before back vowels is 'dark' or velarised ([l^v]). Acoustic data of course do not provide direct information about the articulation of the consonant, so I will

² The data sample contained no Cypriot words where /l/ occurred before front vowels.

use the term ‘velarised’ as a description of the acoustic quality rather than articulatory feature. Higher F2 observed in Thessalian Greek before /i/ may indicate palatalisation.

It appears that both the ‘velarisation’ and the ‘palatalisation’ of /l/ in Thessalian Greek is subject to individual variation. Before /a/, two speakers had a consistently low F2, for others there was much more variation in values. For two speakers who showed consistent velarisation of /l/, mean frequencies of F2 were 1061 Hz and 1138 Hz, which is similar to the values reported for American English ‘dark’ /l/ (see, for example, Huffman 1997). For three Thessalian speakers, in whose data the F2 frequency of /l/ before /a/ was comparable to the Athenian speakers (mean F2 for three Thessalian speakers 1394 Hz, cf. 1490 Hz for Athenian speakers), /l/ before /i/ had significantly higher F2 frequency than the F2 of /l/ of the Athenian speakers (1968 Hz in Thessalian Greek vs. 1690 Hz in Athenian Greek).

As a result, mean difference between F2 of /l/ before /a/ and /i/ for each speaker of Thessalian Greek was 585 Hz compared to 295 Hz in Athenian Greek. Therefore in Thessalian Greek the two variants of /l/ acoustically are further apart than in Athenian Greek, which depending on the speaker can be the result of ‘velarisation’ or ‘palatalisation’ of /l/ or both of these processes.

5. Towards acoustic model of Modern Greek laterals

In the previous sections I suggested that the quality of /l/ in Modern Greek dialects may vary depending on the quality of the following vowel. In this section I will use correlations and multiple linear regression to explore the contribution of F1 and F2 of the adjacent vowels and of the duration of /l/ to the quality of /l/.

The following results are based on a series of multiple linear regressions using the enter method with the F1 or F2 of /l/ as the dependent variables and the following independent variables: F1 and F2 of the preceding and following vowel, duration of /l/ and speaker (recoded into several binary variables).

5.1. Athenian Greek

Multiple regression analysis using the enter method showed that in Athenian Greek about 60% of variation in F1 of /l/ could be explained by F1 of the preceding and following vowel (adjusted R square=0.575, $F_{5,42}=11.354$, $p<0.001$, see Table 4 for coefficients of significant variables). Some variation in F2 frequency of /l/ could be explained by the quality of the following vowel (adjusted R square=0.146, $F_{5,41}=2.569$, $p<0.05$, see Table 4 for coefficients of significant variables).

The effect of the adjacent vowels was mainly due to the difference in the quality of /l/ between front and back vowels. Only in *poli* the frequency of F1 of /l/ was positively correlated with F1 of the following /i/ (Pearson’s $r=0.634$, $p<0.001$) and the preceding /o/ (Pearson’s $r=0.447$, $p<0.05$). There also was positive correlation in F1 between the two vowels (Pearson’s $r=0.500$, $p<0.01$), which may suggest greater degree of coarticulation in this word. The variation in quality of the following /a/ appeared to have no effect on the quality of /l/. There also was no effect of duration.

Table 4: Significant predictor variables in multiple regression analysis in Athenian Greek.

Predictor variable	F1 of /l/		F2 of /l/	
	Beta	p	Beta	p
Duration of /l/				
F1 (preceding vowel)	0.354	$p < 0.01$		
F1 (following vowel)	0.540	$p < 0.001$	-0.477	$p < 0.01$
F2 (preceding vowel)				
F2 (following vowel)			0.325	$p < 0.05$
Adjusted R square	0.575		0.146	

5.2. Thessalian Greek

In Thessalian Greek most variation (83%) in F1 frequency of /l/ could be explained by F1 of adjacent vowels and duration of /l/ (adjusted R square=0.827, $F_{5,90}=92.045$, $p < 0.001$, see Table for coefficients of significant variables). The frequency of F2 of /l/ was also significantly affected by the quality of the adjacent vowels (adjusted R square=0.507, $F_{5,89}=20.370$, $p < 0.001$, see

Table for coefficients of significant variables). I have previously reported inter-speaker variation in F2 in Thessalian Greek. Further analysis showed that differences between speakers account for about 13% of variation: the model that included “speakers” explained 66% of variance (adjusted R square=0.663, $F_{10,84}=19.501$, $p < 0.001$).

As in Athenian Greek, the variation in the quality of /l/ was mainly due to different preceding/following vowel³. In some cases, sub-phonemic variation in F1 or F2 of the preceding or following vowel also had an effect on the quality of /l/. In *kala* the F1 frequency of /l/ was positively correlated with F1 of the preceding (Pearson’s $r=0.633$, $p < 0.001$) and following (Pearson’s $r=0.636$, $p < 0.001$) /a/. As in case of *poli* in Athenian Greek, there also was positive correlation in F1 between the two vowels (Pearson’s $r=0.54$, $p < 0.01$) in *kala* in Thessalian Greek. In *poli* F2 of /l/ was positively correlated with F2 of the following /i/ (Pearson’s $r=0.525$, $p < 0.001$). In *polla* and *kala* the F2 frequency of /l/ was positively correlated with F2 of the previous vowel: Pearson’s $r=0.641$, $p < 0.01$ in *polla* and Pearson’s $r=0.470$, $p < 0.01$ in *kala*.

Table 5: Significant predictor variables in multiple regression analysis in Thessalian Greek.

Predictor variable	F1 of /l/		F2 of /l/	
	Beta	p	Beta	p
Duration of /l/	-0.154	$p < 0.01$		
F1 (preceding vowel)	0.512	$p < 0.001$	-0.228	$p < 0.05$
F1 (following vowel)	0.481	$p < 0.001$	-0.489	$p < 0.001$
F2 (preceding vowel)	-0.109	$p < 0.05$	0.261	$p < 0.01$
F2 (following vowel)			0.255	$p < 0.01$
Adjusted R square	0.827		0.507	

In Thessalian Greek, frequency of F1 decreased with increase in duration: in *poli* and *polla* there was significant correlation between duration and F1 frequency (*poli* Pearson’s $r=-0.345$, $p < 0.05$, *polla* Pearson’s $r=-0.539$, $p < 0.01$).

5.3. Cypriot Greek

In Cypriot Greek some variation in F1 of /l/ could be explained by duration (adjusted R square=0.276, $F_{5,63}=6.185$, $p < 0.001$, see Table 6 for coefficients of significant variables). This reflects the difference between geminates and singletons. There was no further correlation between duration and F1 within each of these categories, that is unlike in

³ It should be remembered that the first vowel in *poli* and *polla* has significantly lower F1 in Thessalian Greek than in Athenian Greek (Loukina 2008; in press).

Thessalian Greek, in Cypriot Greek lower F1 in geminate consonants was not associated with longer duration. There also was a weak effect of F2 of the following vowel.

The effect of adjacent vowels on F2 of /l/ in Cypriot Greek differed between singletons and geminates. In *kala* F2 of /l/ was correlated with the quality of the preceding (F1: Pearson's $r=0.638$, $p<0.05$, F2: Pearson's $r=0.761$, $p<0.01$) and the following vowel (F1: Pearson's $r=0.622$, $p<0.05$, F2: Pearson's $r=0.650$, $p<0.05$). There were no such correlations in *polla*.

Table 6: Significant predictor variables in multiple regression analysis in Cypriot Greek.

Predictor variable	F1 of /l/		F2 of /l/	
	Beta	p	Beta	p
Duration of /l/	-0.346	$p < 0.01$		
F1 (preceding vowel)				
F1 (following vowel)				
F2 (preceding vowel)				
F2 (following vowel)	0.248	$p < 0.05$		
Adjusted R square	0.276		0.066, p = 0.094	

It could be argued that smaller effect of adjacent vowels on the quality of /l/ in Cypriot Greek than in the other two varieties is due to the unbalanced sample: in Cypriot Greek there were no tokens of /l/ before /i/. To test this hypothesis I ran multiple regression analysis on a subset of Thessalian data that only contained tokens of *polla* and *kala*⁴. The results showed that in Thessalian Greek quality of the adjacent vowels accounted for 60% of variation in F1 (adjusted R square=0.602, $F_{5,48}=17.063$, $p<0.001$, see Table for coefficients of significant variables) and 24% of variation in F2 of /l/ (adjusted R square=0.243, $F_{5,48}=4.409$, $p<0.001$, see Table for coefficients of significant variables). This is substantially greater than in Cypriot Greek, which suggests that the observed difference is not an artefact of the sampling method.

Table 7: Significant predictor variables in multiple regression analysis on a subset of data in Thessalian Greek where /l/ only occurred before /a/.

Predictor variable	F1 of /l/		F2 of /l/	
	Beta	p	Beta	p
Duration of /l/	-0.275	$p < 0.01$		
F1 (preceding vowel)	0.830	$p < 0.001$	-0.454	$p < 0.05$
F1 (following vowel)	0.206	$p < 0.05$		
F2 (preceding vowel)	-0.343	$p < 0.01$	0.746	$p < 0.0001$
F2 (following vowel)				
Adjusted R square	0.602		0.243	

6. Discussion

Acoustic analysis of lateral consonants in three varieties of Modern Greek revealed different patterns of variation in duration and quality.

In Cypriot Greek variation in lateral consonants was primarily linked to the contrast between the so-called geminates and singletons. The analysis of spontaneous speech confirmed differences in duration and F1 previously reported for laboratory speech (Payne and Eftychiou 2006, Eftychiou 2008); however, in this data sample there was no difference in F2 between singleton and geminate laterals. There was an overlap between geminates and singletons both in duration and in quality and no evidence for compensatory relations between them. In her EPG study, Eftychiou (2008) found positive

⁴ The number of observations in Athenian Greek was insufficient for the number of predictor variables used in this analysis.

correlation between the amount of contact and duration in Cypriot lateral geminates and raised the question whether the difference in quality between geminates and singletons is the result of temporal difference or a distinct gesture (cf. also Payne 2005; 2006 for the discussion of Italian geminates). In this data there was no correlation between F1 and duration in Cypriot geminates. Noteworthy, in Thessalian Greek, which does not exhibit contrast between geminates and singletons, the variation in F1 frequency of /l/ was correlated with duration: longer consonants had lower F1. The results for Thessalian Greek agree with the temporal explanation: longer consonants allowed for more complete execution of the gesture. The absence of such correlation in Cypriot Greek points towards the independent roles of duration and quality as acoustic correlates of germination in Cypriot Greek.

As expected, the data for Athenian and Thessalian Greek did not show any differences in duration or quality between words spelled with single or double consonants. Despite a lack of contrast between geminates and singletons, laterals in Athenian and Thessalian Greek did not show greater variation in duration or F1 frequency than Cypriot singletons. This shows that the limits of variation are not necessarily determined by the requirement to preserve contrast.

In Athenian and Thessalian Greek variation in lateral consonants was linked to the quality of the following vowel. In both varieties there was significant difference in quality of laterals before /i/ and /a/. The results provided experimental evidence for the existence of [l^v] in Thessalian Greek, which has been often mentioned in impressionistic descriptions. The analysis also revealed significant difference in quality of /l/ before /i/ in Athenian and Thessalian Greek, which may reflect greater palatalisation in Thessalian. In both cases in Thessalian Greek there was individual variation in choice of the variants, but the two variants of /l/ for each individual speaker were consistently further apart in Thessalian Greek than in Athenian Greek. This poses an interesting question: why the velarisation of /l/ is perceived as a more salient dialectal marker than the palatalisation? It may be that it is geographically less widespread than the velarised variant. For example, Kontosopoulos (Kontosopoulos 2001) mentions what can be interpreted as palatalisation of /l/ and /n/ before front vowels "in many parts of Northern Greece, which have not yet been precisely defined by the dialectologists". It may also be that after velarisation had become a stereotype of Northern speech (cf. Labov 1972), it is perceived more readily than palatalisation. The results of this study suggest that the difference between the two variants may be a better measure of comparison between the varieties than acoustic properties of individual sounds, since such difference appears to be more consistent across individual speakers.

It is worth noting that most other Balkan languages once spoken in the same area as Thessalian Greek distinguish between the so-called 'soft' /li/ and 'hard' /l^v/ (see also Jakobson 1931 for a broader discussion of such contrast), including Bulgarian (Tilkov & Boiadzhiev 1981), Macedonian (Minissi *et al.* 1982, Sawicka 2009), Albanian (Kaminskaia 2000), and Aromanian (Lazarou 1986, Kramer 1989, Katsanes & Dinas 1990, Koltsidas 1993).

Experimental phonetic data are only available for Bulgarian. Tilkov and Boiadzhiev (1981) give the following formant frequencies for 'hard' [l^v]: F1=400 Hz, F2=1000 Hz. 'Soft' /li/ has a higher F2 and according to Tilkov & Boiadzhiev the difference in F2 between the two consonants is about 800-1000 Hz (cf. 585 Hz in Thessalian Greek). The F2 of /l/ in my data sample is intermediate to the values given by Tilkov and Boiadzhiev, with Thessalian Greek /l/ before /a/ (F2=1324 Hz) being closer to 'hard' Bulgarian [l^v] than the other two varieties. Thessalian /l/ before /i/ (F2=1839 Hz) is also closer to 'soft' Bulgarian [l] than Athenian Greek. It is likely that the existence of opposition of velarised and palatalised laterals in contact languages contributed to the polarisation of variants of /l/ before back and front vowels in Thessalian Greek.

Sawicka (Sawicka 1997, 2009) suggested that this area of Balkans can be described as characterized by accommodative pronunciation with frequent assimilations and neutralizations. Although one should be cautious when trying to fit all phonetic processes of a given language under single generalization, the results of this study suggest that in Thessalian Greek /l/ is subject to greater coarticulatory effects than in the other two varieties. Quality of adjacent vowels and duration explained 83% of variation in F1 of /l/ in Thessalian Greek, 58% of variation in Athenian Greek and only 28% of variation in Cypriot Greek.

Sproat and Fujimura (1993) suggested that the variation between dark and clear /l/ in English is a result of different phonetic implementation of the same phonological entity depending on the position and duration (cf. also Huffman 1997 for further discussion). They argue that English variation between dark and clear /l/ is continuous and phonetically predictable and there is no need to use distinct phonological units to encode this variation. The results of this study also suggest that the variation in quality /l/ can partially be explained by such factors as the quality of the adjacent vowels and duration. However, the effect of these factors differed across varieties. The effect of the adjacent vowels was very limited in Cypriot Greek, where most of variation was governed by the lexical distinction between geminates and singletons. To the contrary, in Thessalian Greek most of variation in the quality /l/ in this data could be explained by phonetic factors. Further study is needed to establish whether this is true for other positions and contexts. In Athenian Greek the main pattern of variation was similar to the Thessalian Greek, yet the range of variation was significantly smaller and there was less effect of adjacent vowels and no effect of duration. This suggests that such phonetic processes as coarticulation or gestural undershoot may operate to a different degree even in closely related varieties and raises the question of what factors may block or encourage their application.

7. Conclusion

The results reported in this paper are based on a very limited data sample and therefore should be treated with caution. While the conclusions are thus limited, it should be noted that they agree with the results of previous laboratory studies where such exist. This study once again highlighted the non-durational aspects of geminate consonants and provided experimental evidence for the features that until now have been only described on impressionistic basis. It has also revealed new aspects of variation in lateral consonants in Modern Greek dialects.

While a model of phonetic implementation could explain some variation in lateral consonants, the study showed that the rules of phonetic implementation are certainly language-specific (or even dialect-specific in this case). Although the data for each variety consisted of the same lexical items, the patterns of variation in lateral consonants were very different in the three varieties. This once again shows the complexity of interaction between universal physiological principles of speech production and language-specific constraints (cf. also Loukina 2009). There is no doubt that further studies on larger corpora of data from different varieties where lateral consonants would occur in more phonological contexts, different positions and stress conditions will contribute to better understanding not only of differences between Modern Greek dialects and but also general principles of speech production.

Acknowledgments

The results presented in this part are part of a larger study. Financial support for this study was provided by the Oxford University Clarendon Fund, St Cross College and Oxford University Phonetics Lab and is gratefully acknowledged.

References

- Adank P., Smits R. *et al.* (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America* 116(5): 3099-3107.
- Arvaniti A. (1999a). Comparing the phonetics of single and geminate consonants in Cypriot and Standard Greek. Greek linguistics '99. Proceedings of the 4th international conference on Greek linguistics, Nicosia, September 17-19, 1999, Thessaloniki, University Studio Press, 37-44.
- Arvaniti A. (1999b). Cypriot Greek. *Journal of the International Phonetic Association* 29(2): 173-178.
- Arvaniti A. (2001). Cypriot Greek and the phonetics and phonology of Geminate. Proceedings of the First international conference of Modern Greek dialects and linguistic theory (Patras, Greece, Oct. 12-14, 2000), Patras, University of Patras. pp. 19-30.
- Arvaniti A. & Tserdanelis G. (2000). On the phonetics of geminates : evidence from Cypriot Greek. Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP 2000), Beijing, 559-562.
- Eftychiou E. (2008). Lenition processes in Cypriot Greek. PhD Dissertation. Cambridge: University of Cambridge.
- Huffman M. K. (1997). Phonetic variation in intervocalic onset /l/'s in English. *Journal of phonetics* 25(2): 115-141.
- Jakobson R. (1931). К характеристике евразийского языкового союза. Paris, Imp. de Navarre. Reprinted in R. Jakobson. *Selected writings*. The Hague: Mouton & Co, 1962. Vol. 1., 160-201
- Joseph B. & Tserdanelis G. (2003). "Modern Greek." In Variationstypologie. Ein sprachtypologisches Handbuch zu den europäischen Sprachen in Geschichte und Gegenwart, edited by T. Roelcke. Berlin, De Gruyter, pp. 823-836.
- Kaminskaia, L. N. (2000). Вводно-фонетический курс албанского языка. Санкт-Петербург, Санкт-Петербургский Государственный Университет.
- Katsanes N. & Dinas K. (1990). *Γραμματική της κοινής Κουτσοβλάχικης*. Θεσσαλονίκη.
- Koltsidas A. M. (1993). *Κουτσοβλάχοι, οι βλαχόφωνοι Έλληνες : εθνολογική, λαογραφική και γλωσσολογική μελέτη (γραμματική και λεξικό της Κουτσοβλάχικης γλώσσας)*. Θεσσαλονίκη, Εκδοτικός Οίκος Αφων Κυριακίδη
- Kontosopoulos N. G. (2001). *Διάλεκτοι και ιδιώματα της Νέας Ελληνικής*. Αθήνα, Εκδόσεις Γρηγόρη.
- Kramer J. (1989). Rumänisch: Areallinguistik, II: Aromunisch/Les Aires linguistiques, II: Aroumain. In G. Holtus, M. Metzeltin and C. Schmitt. Tübingen, Niemeyer (eds.) *Lexikon der Romanistischen Linguistik (LRL), III: Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart: Rumänisch, Dalmatisch/Istroromanisch, Friaulisch, Ladinisch, Bündnerromanisch/Les Différentes Langues romanes et leurs régions d'implantation de la Renaissance a nos jours: Le Roumain, dalmatico/istroromanzo, friulano, ladino, le romanche*, 423-35.
- Labov W. (1972). *Sociolinguistic patterns*. Philadelphia, University of Pennsylvania Press.
- Ladefoged P. & Maddieson I. (1996). *The sounds of the world's languages*. Oxford, Blackwell.
- Lazarou A. G. (1986). *Βλάχοι : ιστορική φιλολογική μελέτη*. Αθήνα.
- Lobanov B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America* 49(2B): 606-608.
- Loukina A. (2008). *Regional phonetic variation in Modern Greek*. PhD Dissertation. Oxford: Oxford University.
- Loukina A. (2009). Phonetic variation in spontaneous speech. Vowel and consonant reduction in Modern Greek dialects. In E. Payne and O. Parker-Jones (eds.) *Oxford University Working Papers in Phonetics* 12: 36-56. Oxford: University of Oxford.
- Loukina A. (in press). Phonetic variation in regional varieties of Modern Greek: vowel raising. *Proceedings of the 3rd International Conference on Modern Greek Dialects and Linguistic Theory*, Nicosia, 14-16 June 2007.
- Minissi N. *et al.* (1982). *Fonetika na makedonskiot jazik = The phonetics of macedonian = La fonetica del macedone*. Napoli, Bibliopolis.
- Papazachariou D. (2003). Οι πραγματώσεις της φωνολογικής μονάδας /L/ της Πατρικής διαλέκτου. *Proceedings of the sixth International Conference of Greek linguistics*, Rethymno, 18-21 September 2003, Rethymno, Linguistics Lab. pp. CD-Rom (ISBN 960-88268-0-2).
- Payne E. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2): 153-181.
- Payne E. (2006). Non-durational indices in Italian geminate consonants. *Journal of the International Phonetic Association* 36(1): 83-95.

- Payne E. & E. Eftychiou (2006). Prosodic Shaping of Consonant Gemination in Cypriot Greek. *Phonetica* 63(2-3): 175-198.
- Peterson G. E. & I. Lehiste (1960). Duration of Syllable Nuclei in English. *The Journal of the Acoustical Society of America* 32(6): 693-703.
- Sawicka I. (1997). *The Balkan Sprachbund in the light of phonetic features*. Warszawa, Wydawnictwo Energeia.
- Sawicka I. (2009). The phonetics of Macedonian. *Language Typology and Universals* 61(2): 147-152.
- Sproat, R. & O. Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291 - 311.
- Tilkov D. & Boiadzhiev T. (1981). *Българска фонетика*. София, Наука и изкуство.
- Tserdanelis G. & Arvaniti A. (1999). The acoustic characteristics of geminate consonants in Cypriot Greek. Greek linguistics '99. *Proceedings of the 4th international conference on Greek linguistics*, Nicosia, September 17-19, 1999, Thessaloniki, University Studio Press, 29-36.

Syntactic micro-variation in Pontic: Dative constructions¹

DIMITRIS MICHELIOUDAKIS
University of Cambridge
dm419@cam.ac.uk

IOANNA SITARIDOU
University of Cambridge, Queens' College
is269@cam.ac.uk

1. Introduction

This paper is a first attempt at a syntactic analysis of dative constructions in Pontic Greek (PG) (see also Drettas 1997), quite an understudied syntactic phenomenon, and an inadequately explored area in the study of Greek dialects, in general (but see Manolessou & Beis 2006 for a general overview).

Drawing data from three different varieties of Pontic Greek namely, Romeyka of Of (ROf), Romeyka of Sürmene (RSür)² –both spoken in Turkey– and Pontic Greek (PG) as spoken in Thessaloniki, we set out to explore all the possible patterns in the syntax of the substitutes of the Ancient Greek (AG) dative. In doing so, we relate them to some more general properties of double-object constructions and dative alternations, whilst also trying to specify the status of the PG ‘datives’ with regards to the ‘inherent’ vs. ‘structural’ distinction.

It is claimed that: (a) Romeyka (both Of and Sürmene varieties) lacks dative alternations despite having the double DP frame for ditransitives; (b) The underlying hierarchical relations in Romeyka are the reverse from what we find in PG; (c) PG behaves syntactically on a par with Standard Modern Greek (SMG) despite the differences in the morphological realisation of the DPs (and which is almost identical in all Pontic varieties namely, mACC for both arguments); (d) ROf and RSür behave identically with the exception of the benefactives where we find more intense microvariation; (e) In all three Pontic varieties clitic movement of the dative arguments –which is otherwise obligatory in SMG– is not required in unaccusatives and passives thus indicating that the Case feature in these varieties is such that it does not cause any minimality effects, i.e., non-quirky, purely inherent.

The paper is organised as follows. In section 2 we discuss the structural representation of dative arguments in SMG in order to establish a comparative platform for the Pontic data to be examined in section 3; in particular, we discuss the prototypical ditransitive

¹ We are extremely grateful to our informants: T. for ROf, Hakan Özkan for RSür, and LEMONIA Tsakiridou for PG. All errors are our own. Dimitris Michelioudakis wishes to thank the Greek State Scholarships Foundation (IKY) and the Alexander Onassis Foundation for funding his graduate research.

² This article forms part of a larger project on the syntax of the Romeyka (Hellenic) varieties of Pontus “Contact, continuity and change: The syntax of the Romeyka varieties in Pontus” (PI: Ioanna Sitaridou, <http://people.pwf.cam.ac.uk/is269/research-projects.html>). For a general view on microvariation in the Pontic varieties see Kaltsa and Sitaridou (2009). For other syntactic phenomena in Pontic varieties see Sitaridou (2009a), Kaltsa & Sitaridou (this volume). On the methodology, taxonomy and language use of the Of variety in Pontus see Sitaridou (2009b). From a glossonymic perspective, we use the term ‘Romeyka varieties of Pontus’ to refer to what is previously known as ‘Muslim Pontic’ (as in Mackridge 1987). When further specification is required –‘Romeyka varieties of Pontus’ is an umbrella term after all (cf. Sitaridou 2009b)– we further specify it as ‘Romeyka varieties of Of’, ‘Romeyka varieties of Sürmene’. The methodology we used entailed oral interviews comprising structured questionnaires. The only variants we controlled for –and the only ones we think are relevant for this phenomenon/varieties– are: (i) geographical location of the speaker; and (ii) the degree of exposure to either Standard Modern Greek (SMG) or Turkish.

constructions in section 3.1, benefactives in 3.2 and experiencers in 3.3 whereas in 3.4 we summarise our findings. Finally, in section 4 we present our concluding remarks with regard to mCase and the abstract ‘dative’ features.

2. The structural representation of dative arguments in Standard Modern Greek

SMG has two structural representations for ditransitives (cf. Anagnostopoulou 2003), as shown in (1):

- (1) $DP_{IO} \gg DP_{DO}$ (where ‘ \gg ’ means asymmetric c-command)
- a. O Kesaras eðikse [tu kathe ðiikiti]_i [tin eparxia tu]_i (sto(n) xarti) (SMG)
 Caesar showed.3SG [the.GEN every/each governor.GEN] [the province.ACC his]
 (on the map)
 ‘Caesar showed every/each governor his province (on the map)’
- b. ?*O Kesaras eðikse [tu ðiikiti tis]_i [kathe eparxia]_i
 Caesar showed.3SG [the.GEN governor.GEN his] [every/each province.ACC]
 ‘Caesar showed its governor every/each province’
- c. ??O Kesaras eðikse [kathe eparxia]_i [tu ðiikiti tis]_i
 Caesar showed.3SG [every province.ACC] [the.GEN governor.GEN her]
 ‘Caesar showed every province its governor’

From (1) we conclude that $IO \gg DO$ is the underlying representation in DP_{IO} - DP_{DO} constructions (as well as in Benef-DP-Acc-DP constructions).

In the prepositional constructions however, the reverse pattern is found, as shown in (2):

- (2) $DP_{DO} \gg PP_{IO}$
- a. O Kesaras eðikse [tin kathe eparxia]_i [sto(n) ðiikiti tis]_i (SMG)
 Caesar showed [the every province.ACC] [to+the governor her]
 ‘Caesar showed every province its governor’
- b. ?*O Kesaras eðikse [tin eparxia tu]_i [se kathe ðiikiti]_i
 Caesar showed.3SG [the province.ACC his] [to every governor.ACC]
 ‘Caesar showed his province to every governor’

$PP_{IO} \gg DP_{DO}$ may optionally appear in the promoted position (the one that asymmetrically c-commands the DO) in goal-ditransitives (3a) and obligatorily in benefactives (3b, 3b’):

- (3) $PP_{IO} \gg DP_{DO}$
- a. O Kesaras eðikse [se kathe ðiikiti]_i [tin eparxia tu]_i (SMG)
 Caesar showed.3SG [to every governor.ACC] [the province.ACC his]
 ‘Caesar showed to every governor his province’
- b. O Kesaras sxeðiase [se kathe ðiikiti]_i [ena sxeðiagrama tis eparxias tu]_i
 Caesar drew.3SG [to every governor.ACC] [a map/diagram.ACC the.GEN province.GEN his]
 ‘Caesar drew to every governor a map/diagram of his province’
- b’. ?* O Kesaras sxeðiase [ena sxeðiagrama [kathe eparxias]_i] [ston ðiikiti tis]_i
 Caesar drew.3SG a diagram.ACC every province.GEN to+the governor.ACC his
 Caesar drew [a diagram of every province] [to its governor]

Tables 1 summarises the c-command relations of IO and DO found in SMG whereas Table 2 does the same for benefactives:

Table 1: *c-command relations in goal ditransitives (SMG)*

	IO>>DO	DO>>IO
DPgen	✓	*
<i>se</i> 'to'-PP	✓	✓

Table 2: *c-command relations in benefactives (SMG)*

	Benef>>DO	DO>>Benef
DPgen	✓	*
<i>se</i> 'to'-PP ³	✓	*

3. Microvariation in Pontic dative constructions

The underlying order IO>>DO when both arguments are DPs does not hold across all Greek varieties, as will be discussed below. The standard pattern is (partially) replicated only in PG.

3.1. Recipients/Goals

In section 3.1.1 we discuss the Romeyka varieties (Of and Sürmene) whereas in 3.1.2 we discuss the PG variety.

3.1.1. Romeyka varieties of Pontus (Of and Sürmene)

- IO DPs are accusative and do not alternate with PPs:

- (2) a. To peði eđotfe fai ton ađelfo / *son ađelfo (RSür)
 the child gave-3SG food the brother.ACC / *to-the brother
 'The child gave food to the brother'
 b. To peði eđose fai ston ađelfo (SMG)
 the child gave-3SG food to+the brother
 'The child gave food to the brother'

- Both surface orders (IO-DO and DO-IO) are licit:

- (3) a. To peði eđotfe fai ton ađelfo / ton ađelfo fai (RSür)
 the kid gave.3S food the brother / the brother food
 'The kid gave food to the brother'
 b. Eyo eđoka ton Mehmeti ena kitap / ena kitap ton Mehmeti (ROf)
 I gave.1S the Mehmet a letter (?) / a letter the Mehmet
 'I gave Mehmet a letter'

- PP-realisation is restricted to purely locative uses:

- (4) Epije so kulin (ROf)
 went.3SG to-the school
 'He went to the school'

- Barss & Lasnik's (1986) diagnostics for *c-command* indicate that DP_{DO} asymmetrically *c-commands* DP_{IO}:

- (i) Weak Crossover Effects:

³ These benefactive PPs may optionally be introduced with the preposition *ja* 'for' in SMG. In this case the benefactive PPs seem to occupy an adjunct position *c-commanding*, but otherwise unable to bind the DO.

- (5) a. Pion zon ekloses ton tʃopanonat? (RSür)
 which animal sent.2S the shepherd-its?
 ‘Which animal did you send to its shepherd?’
- b. *Tinan tʃopan(i) ekloses to zonat? (RSür)
 which shepherd sent.2S the animal-his?
 ‘Which shepherd did you send his animal to?’
- (ii) Superiority effects (Romeyka has multiple *wh*-fronting which always obeys superiority, cf. the subject-object asymmetry in (6)):
- (6) a. Pion ospit tinan eðikses? (ROf)
 which house whom showed.2SG?
 b. *Tinan pion ospit eðikses? (ROf)
 Whom which house showed.2SG
- (7) a. Pios tinan iðen? (ROf)
 who.NOM whom.ACC saw.3SG
 ‘Who saw whom?’
 b. *Tinan pios iðen? (ROf)
 whom.ACC who.NOM saw.3SG
 ‘Whom did who see?’
- (iii) Quantifier variable binding:
- (8) ta yarðelæ xore xore eðiksa tʃi maylimis’atun (ROf)
 the children every every showed.1SG the teachers-their
 ‘I showed all the children, one by one, to their teachers (each child to its own teacher)’
 *‘I showed every child his/her teacher’

Table 3 summarises the c-command relations of IO and DO found in Romeyka:

Table 3: *c-command relations in goal ditransitives (ROf, RSür)*

	IO>>DO	DO>>IO
DPacc	*	✓
se ‘to’-PP	*	*

This is quite an important finding, as it seems that underlying DO>>IO in the double DP construction is not non-existent or unique to German, for which the same diagnostics lead to the same conclusion (as in Müller 1995, 1999 and McGinnis 1999). In fact, the situation seems to be the same in some historical varieties of Greek, notably Medieval Cypriot Greek (for a discussion of Medieval Cypriot Greek double object constructions see Michelioudakis 2009). This constitutes a serious challenge for the validity of the cross-linguistic generalisation that IOs merge higher than DOs. Furthermore, the observation that the IO is asymmetrically c-commanded by the DO also ties in well with the fact that direct passives are entirely unproblematic in such languages, since the low position of the IO cannot cause any locality effects.

- Direct Passives: In passives, the theme Agrees with T and becomes nominative (and, possibly, moves to a subject-position), without the requirement that the dative argument cliticise (9a, 9b), contrary to SMG (9e) and PG (9c, 9d), which patterns with SMG in this respect. Therefore, the IO DP in these constructions does

not cause any minimality effect in the relation between T and the theme, either because (i) it is not an active goal, i.e. it does not have any uninterpretable (Case) feature, or (ii) because it simply does not intervene structurally, by being lower than DO, as we argued above, or actually because of both (i) and (ii), as we will argue in section 4.

- (9) a. I para tin Aife eðoste (RSür)
 the money.NOM the Ayshe.ACC was-given.3S
 ‘The money was given (to) Ayshe’
 b. To harti eyrafte tin Aife (RSür)
 the letter.NOM was-written the Ayshe.ACC
 ‘The letter was written for (/sent to) Ayshe’

On the contrary, in varieties with hierarchically high IOs (which probably also carry an active Case feature, see section 4), direct passives are impossible unless the IO undergoes clitic-movement:

- c. *Para eðothen tin Anastan (PG)
 money.NOM was-given.3SG the Anasta.ACC
 ‘The money was given (to) Anasta’
 d. ??Tin Anastan eyraften-aten to gramana (PG)
 the Anasta.ACC was-written-Cl.ACC.3SG.FEM the letter.NOM
 e. ta lefta *(tis) epistrafikan tis Marias (SMG)
 the money.NOM her.GEN were-given the.GEN Maria.GEN
 ‘The money was returned to Maria’

- Clitic clusters: Prima facie, it looks like Romeyka may have clitic clusters (10a-c).

- (10) a. Eðiksen aton(a) (ROf)
 showed.3SG him
 ‘(S)he showed him’
 b. I Aife eðotfen aton ena pita (ROf)
 the Ayshe gave.3SG him a pie
 ‘Ayshe gave him a pie’
 c. Eðiksane-me aton(a) (RSür)
 showed.3PL me him
 ‘They showed him to me’

However, a closer inspection reveals that in Romeyka, unlike PG, the 3SG personal pronoun /ato(n)(a)/ does not have clitic-like properties (10d-g):

- d. Eðotfen-eme o Mehmet ato(n) (ROf)
 gave.3SG Cl.1SG.ACC the Mehmet.NOM him/it.ACC
 ‘Mehmet gave me this/it’
 e. Eðiksane to Mehmet atona (RSür)
 showed.3PL the Mehmet him
 ‘(?)They showed Mehmet to him’
 f. O Mehmeyis adona etfino fanerose (RSür)
 the Mehmet.NOM him.ACC this.ACC showed.3SG
 ‘Mehmet showed this to him’
 g. Eðotfen eme o Mehmet aton (ROf)
 gave.3SG me the Mehmet.NOM him/it

‘Mehmet gave him/it to me’

Also, interestingly, the corresponding clitic /æ/ cannot combine with any other clitic in any person (11):

- (11) a. O Mehmetis emenan eđotʃen-æ (ROf)
 the Mehmet.NOM me.ACC gave.3SG-Cl.3SG.ACC
 ‘Mehmet gave it to me’
 b. *O Mehmetis eđotʃe-m(e)-æ (unattested in ROf, OK in PG)
 the Mehmet gave.3SG-Cl.1SG.ACC-Cl.3SG.ACC
 ‘Mehmet gave it to me’

- Person-Case effects (restrictions on the person specification of DO in the presence of a dative, see Bonet 1991):

- (12) a. Eđiksane me/emenan atona (RSür)
 showed.3PL Cl.1SG.ACC/me.ACC him.ACC
 b. Eđiksan(e) æ /aton(a) emenan (RSur/ROf)
 showed.3PL Cl.3SG.ACC/him.ACC me.ACC
 ‘They showed him to me / *They showed me to him’
- (13) a. Eđiksane-m’ ese / *eđiksane-s’ eme (RSür)
 showed.3PL-Cl.1SG.ACC you.ACC / showed.3PL-Cl.2SG me.ACC
 b. Atos esena emen eđikse (ROf)
 He you.ACC me.ACC showed.3SG

Interestingly enough, Person-Case effects are not absent from Romeyka, despite the lack of clitic clusters. Combinations of strong pronouns, or of clitics and strong pronouns (12), are subject to the PCC, though a weaker version of it: the sequences of a 1st person clitic and a 2nd person pronoun (cf. 13) are acceptable for most of the speakers, and surprisingly the same pattern (as in 13a-13b) is attested in some Pontic varieties of Northern Greece (Chatzikiyriakidis, 2010). Recall that SMG has the strong version of the PCC (13c). It is an open question if the examples in (13) (the grammatical ones) can mean both ‘they showed you to me’ and ‘they showed me to you’.

- c. *Mu se eđiksan
 Cl.1SG.GEN Cl.2SG.ACC showed.3PL
 ‘They showed you to me’

It is worth noting that the equivalent of (12b) in SMG (12), with an IO-clitic and a strong pronominal 1st person DO, would be perfectly grammatical on the reading ‘They showed *me* to him’; however, the use of the strong pronoun in this context is inherently emphatic (as e.g. in Italian, see Bianchi 2006), while in Romeyka this is the unmarked option (see Michelioudakis (to appear) for further details).

- (14) Tu eđiksan *emena* (SMG)
 Cl.3SG.GEN.MASC showed.3PL me.ACC
 ‘They showed *me* to him’

3.1.2. Pontic varieties of Northern Greece (PG)

PG patterns with SMG with respect to the hierarchical/c-command relations between IO and DO (15-17) and the availability of direct passives (see 9c-9d above). Like Romeyka, PG employs morphological accusative DPs for indirect objects, but those also alternate

with PPs (16b, 17b) (*se*-PPs also appear in constructions with underlying IO>>DO, but when *wh*-fronted they can only be bare accusatives; it might be the case that *se* has become more of a Case marker, especially in the fusional determiner *son/sin/so* [*se+ton/tin/to*]='to+the'. Also, PG arguably has clitic clusters, with an IO-DO order.

- To test for the hierarchical/c-command relations between IO and DO we employ Barss and Lasnik's (1986) diagnostics:

(i) Superiority effects:

- (15) a. *Tinan pion ospit eðiksisēs?* (PG)
 whom.ACC which house.ACC showed.2SG
 b. *?*Pion ospit tinan eðiksisēs?*
 which house.ACC whom.ACC showed.2SG
 'Which house did you show to whom?'

(ii) WCO:

- (16) a. *Tinan eðiksisēs t' ospitn-at?* (PG)
 whom.ACC showed.2SG the house.ACC-his
 '(to) whom did you show his house?'
 b. *Pion ospit eðiksisēs son kyrn-at / *ton kyrn-at?*
 which house.ACC showed.2SG to-the owner.ACC-its / the owner.ACC-its
 'Which house did you show to his owner?'

(iii) Quantifier variable binding:

- (17) a. [*Enan enan ta peðia*]_i eðiksan ton ðeskalon-at_i (PG)
 one one the children.ACC showed.3PL the teacher.ACC-its
 'They showed every child (one by one) his/her teacher'
 b. [*Enan enan ta peðia*]_i eðiksan-ato_i son ðeskalon-at_i / *ton ðeskalon-at_i
 one one the children showed.3PL-Cl.3S.ACC to-the teacher-its/the teacher-its
 'They showed every child to his/her teacher'

Table 4 summarises the c-command relations of IO and DO found in PG where we observe the same pattern as in SMG.

Table 4: *c-command relations in goal ditransitives (PG)*

	IO>>DO	DO>>IO
DP _{acc}	✓	*
<i>se</i> 'to'-PP	✓	✓

3.2. Benefactives

As in the case of genuine (goal) ditransitives, both surface/linear orders (IO-DO and DO-IO) are attested in benefactives too in (almost) all varieties (18). Additionally, benefactives may alternate with PPs headed by *ðæ* 'for' (in RO_f) or *ja* 'for' (RS_{ür} and PG), the use of which seems obligatory in direct passives (19). However, there is a dispreference for the DP_{DO}>DP_{Benef} structure, especially when the beneficiary is not the potential/intended recipient –let us call them 'on behalf of/for someone's sake'-benefactives.

Although our data still do not give us conclusive indications, a first approximation about the c-command relations of benefactives would be to categorise them on the basis of two main factors: (i) The distinction mentioned above namely, between

‘(potential/intended) recipient’ benefactives and ‘on behalf of’-benefactives. This distinction is relevant for ROf and PG, where beneficiaries may appear as adjuncts c-commanding [V DO], in which case they can neither bind the DO (because they are not in an A-position) nor be bound by it (since they do not bind it), which is why the Quantifier Variable diagnostic is not applicable; ‘recipient’-benefactives may either merge as adjuncts or in the low position (associated with goals/recipients) c-commanded by DO (20b, 20c), whereas ‘on behalf of’-benefactives can only merge as adjuncts (21a); (ii) The availability of High Applicatives (Pylkkänen 2002): In RSür, all benefactives are being reanalyzed as high applicative arguments c-commanding DO and not vice-versa (20a, 21b). This may also entail some change in the character/content of its [Case] feature (see section 4), i.e. the emergence of a ‘quirky’ inherent Case feature as in SMG, which is able to cause intervention effects; this would explain the unavailability of direct passives with benefactives in this variety (19b) as the impossibility of raising DO to T across the dative; direct passives are ruled out in ROf (19a) anyway, even when the dative is a genuine (low) IO, probably for independent reasons (there is a number of Greek varieties that avoid passivisation after all).

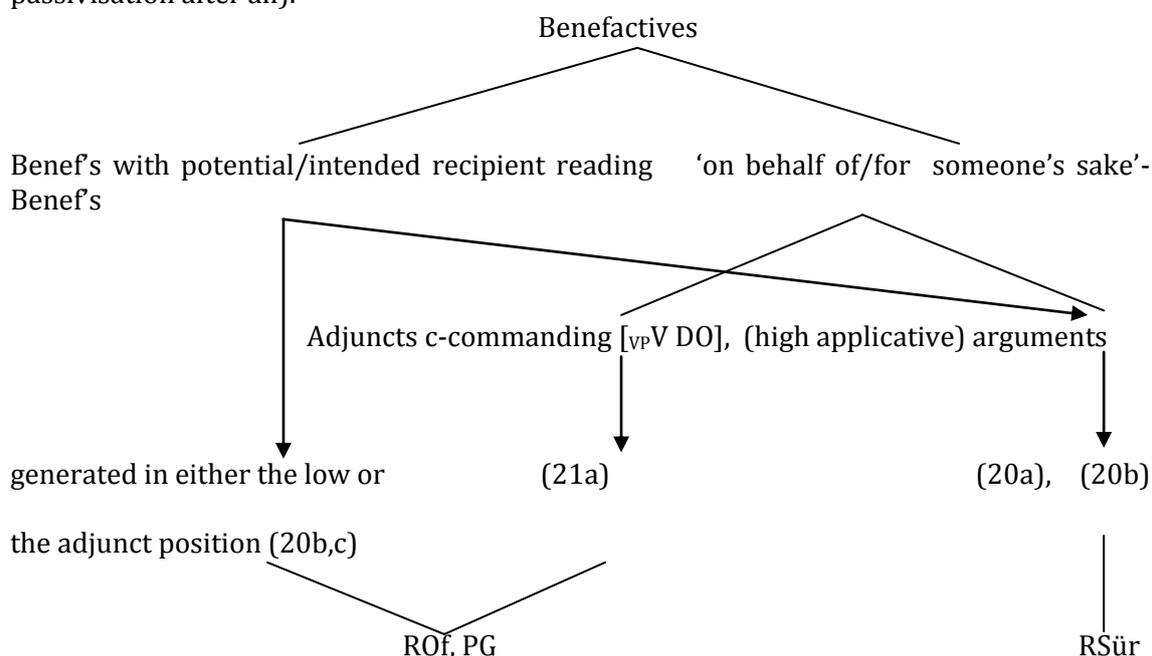


Figure 12: *Benefactives in different Pontic varieties*

- (18) a. Aife epitʃe to Mehmet pide / pide to Mehmet (RSür)
 Ayshe made.3SG the Mehmet.ACC pie.ACC / pie.ACC the Mehmet.ACC
 ‘Ayshe baked Mehmet a pie’
- b. I Aife epitʃen aton enan pita / ?enan pita aton (ROf)
 the Ayshe.NOM made.3SG him.ACC a pie.ACC / a pie.ACC him.ACC
 ‘Ayshe baked him a pie’
- c. I Anasta epiken pitan ton Lefteri / ?ton Lefteri pitan (PG)
 The Anasta.NOM made.3S pie.ACC the Lefteris.ACC/the Lefteris.ACC pie.ACC
 ‘Anasta baked Lefteris a pie’
- (19) a. i pita *(ðæ) ton mehmet epsethen (ROf)
 the pie.NOM for the Mehmet.ACC was-baked
 ‘the pie was baked for Mehmet’
- b. Avuto i pasta *(ja) to Mehmet epsethe. (RSür)
 this pie.NOM for the Mehmet.ACC was baked.3SG
 ‘This pie was baked for Mehmet’
- c. I pita emairefte son Lefteri (PG)

this pie.NOM was-cooked.3SG to-the Mehmet.ACC
 'This pie was baked for Lefteris'

- (20) a. (Ja) tinan d' epitʃe? / *Do tinan epitʃen? (RSür)
 whom.ACC what.ACC made.3SG / What.ACC whom.ACC made.3SG
 b. Tinan tohna epitʃen? / tohna tinan epitʃen? (ROf)
 whom.ACC what.ACC made.3SG / what.ACC. whom.ACC made.3SG
 c. Tinan ti epiken? / Ti tinan epiken? (PG)
 whom.ACC what.ACC made.3SG / What.ACC whom.ACC made.3SG
 'What did she make for whom?'
- (21) a. Tinan tʃopan_i efaises to zon-at_i? / ?Pion zon efaises ton tʃopanin-at? (PG/ROf)
 Which shepherd fed.2SG his animal/which animal fed.3SG his shepherd.ACC
 b. (Ja) tina tʃopano ta provatat efaises? / *Pio provat efaises ton tʃopan-at? (RSür)
 (for) which shepherd the sheep-his fed.2SG/which sheep fed.2SG the shepherd-its
 'For which shepherd did you feed his sheep? / Which sheep did you feed for
 its/their shepherd?'
- (22) a. *O Mehmet etreksen / jelase tin Aiʃe (ROf, PG)
 the Mehmet ran.3SG / smiled.3SG the Ayshe.ACC
 'Mehmet ran for Ayshe / smiled for/at Ayshe'
 b. O Janis ?*(tis) etrekse / ?*(tis) hamojelase tis Marias (SMG)
 the John Cl.GEN.3SG.F ran.3SG/Cl.GEN.3SG.F smiled.3SG the Mary.GEN
 John ran for Mary / smiled for/at Mary
 c. O Mehmetis sin Aiʃe / *tin Aiʃe merea etrehse (RSür)
 the Mehmet.NOM to-the Ayshe.ACC / the Ayshe.ACC merea?? Ran.3SG
 'Mehmet ran to / *for Ayshe'
 d. O Mehmetis tin Aiʃe examojelase (RSür)
 the Mehmet.NOM the Ayshe.ACC smiled
 'Mehmet smiled for/at Ayshe'

Table 5 summarises the c-command relations of Beneficiary and DO found in all varieties of Pontic:

Table 5: *c-command relations in benefactives (all varieties of Pontic)*

	Benef>>DO	DO>>Benef
DPacc	✓ (in all varieties, esp. with non-recipients)	* (RSur), ?/%✓ (ROf, PG)
<i>se</i> 'to'-PP	* (RSur, ROf), ✓ (PG)	* (RSur, ROf), no PG data
<i>ja/ðæ</i> 'for'-PP	✓ (RSur, ROf)	✓ (RSur, ROf, only with potential recipients)

3.3. Unaccusative with datives/experiencers

The use of Class III (*piacere*-type) psych-predicates is rather limited in Pontic, especially in the Romeyka varieties. To the extent that they are used, at least in PG and ROf, they most probably involve the same thematic hierarchy as their equivalents in SMG, Italian etc. (for instance, they allow for backward binding of the nominative theme by the dative experiencer (23)).

- Class III experiencers allow backward binding:

(23) O eaftonats ki ares sin Aiʃe (PG)
 The self-her.NOM not appeal to-the Ayshe

‘Ayshe does not like herself’

However, in Romeyka, as the example (24) from ROF indicates, T-Agree with the theme across the experiencer DP is unproblematic, without any blocking effects or the requirement that the experiencer cliticise (as in SMG). Also, again unlike SMG, which allows PP- and DP-experiencers of such predicates to have subject-like behaviour, quirky experiencer subjects are clearly not possible in Romeyka (see 25 from ROF).

- Class III experiencers do not cause intervention effects (in theme raising) in Romeyka:

- (24) a. I patshi to Hosni aresi (ROF)
the girl the Hosni appeals-to.3S
‘The girl appeals to Hosni’
b. I musiki ?*(tu) aresi tu Jani (SMG)
the music Cl.GEN.3SG.MASC appeal.3SG the John.GEN
‘John likes music’

- ‘Dative’ experiencers do not exhibit subject-like behaviour in Romeyka, unlike SMG:

- (25) O Abdulah_i tin Aiʃen eghapenen (/ *Ton Abdulah_i i Aiʃe aresen), (ROF)
ama pro_i tin Eminen epiren The Abdulah.NOM the Ayshe.ACC loved.3SG /
the Abdulah.ACC the Ayshe.NOM appealed-to.3SG, but pro the Emine.ACC married.3SG
‘Abdulah liked Ayshe, but he married Emine’

cf. SMG, in which dative experiencers can be co-ordinated with nominative null subjects (the diagnostic in (26) is copied from Anagnostopoulou (1999), and the ungrammaticality of the co-indexed *aftos* ‘he’, which is a demonstrative pronoun and causes a Principle C violation, suggests that the dative is in an A-position):

- (26) O Janis_i aghapuse tin Eleni (/ Tu Jani_i *(tu_i) arese i Eleni), ala pro_i / (SMG)
*aftos_i pandreftike ti Maria
the John.NOM loved.3SG the Helen.ACC (the John.GEN Cl.GEN.3SG.MASC appealed-
to.3SG the Helen.NOM) but pro married the Mary.ACC
‘John loved/liked Helen, but he married Mary’

It is striking that PG is attrited by SMG to such an extent that it has lost morphologically accusative Class III experiencers (27); instead, it has genitive and PP ‘dative’ experiencers just like SMG does.

- (27) Ti Abdulah aresen i Aiʃe ebron aso (/ atos) na inekiz me tin Emine (PG)
the.GEN Abdulah appealed.3SG the.NOM Ayshe.NOM before na married.3SG with
the.ACC Emine
‘Abdulah liked Ayshe before he married Emine’

- (28) Ti Mexhmet ke ti Aifēs aresi o enas (s)ton alon (PG)
the Mehmet.GEN and the Ayshe.GEN appeal-to.3PL the one (to-)the other
‘Mehmet and Ayshe like each other’

Interestingly, despite the morphological influence, unlike SMG, there is no blocking effect by the genitive experiencer and no need for cliticisation in the PG examples.

- Romeyka allows (morphologically accusative) goal DPs with motion unaccusative predicates, and again no intervention effects arise in T's Agree with the theme.

(29) To xarti to Meme epiġe (RSür)
 the paper.NOM the Mehmet.ACC went.3SG
 'The letter came/arrived/went (to) Mehmet'

3.4. A comparative table of the findings in all three varieties

Table 6 summarises all our findings so far including information about ethic datives and *wh*-fronting which although not treated here there seem to correlate with the properties discussed in this paper.

Table 6: *Comparative findings across Pontic*

Property	RSür	ROf	Of attrited (Turkish influence)	Of attrited (Greek influence)	PG	
DO _{acc} -IO _{acc} (surface order)	Yes	yes	yes (V-final)	no *(PP)	%	
IO _{acc} -DO _{acc} (surface order)	Yes	yes	yes (V-final)	no *(PP)	Yes	
Locative PPs	Yes	yes	yes	Yes	Yes	
Argumental PPs	No	no	no	Yes	Yes	
Direct Passives	Yes	no	no	no (only benefactive PPs)	No	
Indirect Passives	??	No	no	No	??	
Benefactive PPs	<i>son</i>	No	no	no	Yes	yes
	Other (<i>ja, ðæ</i>)	Yes	yes	yes	No	??
Benefactive Acc	Yes	yes	yes	No	Yes	
Benefactive _{acc} -DO	Yes	yes	yes (V-final)	Yes	(?*(P)DP)	
DO-Benefactive _{acc}	Yes	no	no *(PP)	no *(PP)	no *(PP)	
Ethical Dative	No	no	no	Yes	%	
Barss & Lasnik's tests (suggesting DO>>IO _{acc})	Yes	yes			only with PP-IOs	
Barss & Lasnik's tests (suggesting IO _{acc} >>DO)	No	no			Yes	
Barss & Lasnik's tests (suggesting DO>>Benef _{acc})	No	yes			No	
Barss & Lasnik's tests (suggesting IO>>Benef _{acc})	Yes	yes			Yes	
CD with DO	yes (limited)	no	no	No	yes (limited)	
CD with IO	No	no	no	No	yes (limited)	
Clitic clusters	No	no	no	No	yes(?)	
multiple <i>wh</i> -fronting	Yes	yes	yes	Yes	Yes	
PCC	weak	weak	no(?)	Weak	yes/weak	

Experiencers (Cl.)	(non-obl.)	?	acc	acc	?	Gen
-----------------------	------------	---	-----	-----	---	-----

4. Concluding remarks

The syntactic behaviour and distribution of IOs and experiencers in RO_f and RS_{ür} (and also, surprisingly, experiencers in PG) point towards the hypothesis that abstract ‘dative’ in these cases is a completely interpretable (hence, inactive for Agree purposes) inherent Case feature. This leads us to postulating the typology in Table 7:

Table 7: *m-Case and the structural vs. inherent distinction (all varieties of Pontic)*

+/-Quirky	Quirky	Non-quirky
mACC	PG	RO _f , RS _{ür}
mGEN	SMG	? (MedCG)

Apart from the lack of any intervention effects, quirky subjects etc. with datives in these varieties, the assumption about a fully interpretable, non-quirky Case feature is also made necessary by the fact that only such a feature would survive in the low IO position in the [_{v*P} EA v-V [_{VP} DO <V> IO]] structure that we posit for ditransitives in Romeyka; otherwise, it could not Agree with a phi-probe and get deleted because of the intervention of the DO by virtue of being in a higher position.

In PG, as in SMG, ditransitives (and ‘recipient’ benefactives), which allow for dative shift, probably involve a more articulate structure (essentially in the spirit of Larson 1988), such as [_{v*P} EA v* [_{AppIP} IO Appl [_{v2P} v2 [DO V <IO>]]]], which includes 2 phi-probes, and which may be a necessary condition for clitic doubling. Dative arguments in such constructions probably involve a quirky inherent Case feature, partially unvalued/uninterpretable, which renders the ‘dative’ active for Agree/Move. This Case feature, by having an uninterpretable/‘structural’ part, forces them to occupy (by internal Merge) the edge of an applicative head, where they can Agree with some phi-probe, either v* (in ditransitives), or T when datives with the same feature appear in passives/unaccusatives (see SMG Class III experiencers, which cause intervention effects in T-Agree and have optional subject-like behaviour).

Moreover, pure inherent Case (iCase) causes no minimality effects (phi-probes look for [uCase] in constructions such as raising and unaccusatives), whilst (even partially) uninterpretable Case features (quirky Case) do not. Valuation of quirky Case takes place prior to T’s (further) probing, so this is an instance of defective intervention (Chomsky 2001); (obligatory) dative clitic-movement (in SMG) obviates this defective intervention effect, since the new head of the dative’s chain, i.e. the clitic, is outside T’s Agree domain. An interesting case of micro-variation in this respect is that in SMG, as said above, Class III experiencers have quirky properties (e.g., intervention effects in T-Agree, optional subject-like behaviour), but these are entirely absent from PG, which may mean that quirky inherent Case in SMG spread from goal/benefactive arguments to experiencers, a change which may have not yet taken place in PG. Finally, the apparent availability of high applicatives in RS_{ür} may be a first step for the emergence of quirky inherent Case in this variety too.

References

- Anagnostopoulou E. (1999). ‘On Experiencers’. In A. Alexiadou, G. Horrocks, and M. Stavrou (eds.) *Studies in Greek Syntax*: 67-93. Dordrecht, Holland: Kluwer Academic Publishers
- Anagnostopoulou E. (2003). *The Syntax of Ditransitives. Evidence from Clitics*. Berlin-New York: Mouton de Gruyter.
- Barss, A. & Lasnik H. (1986). A Note on Anaphora and Double Objects. *Linguistic Inquiry* 17: 347-354

- Bianchi V. (2006). On the syntax of personal arguments. *Lingua* 116: 2023-2067.
- Chatzikyriakidis S. (2010). Clitic Clusters in Pontic. Poster presented at ‘The Greek language in Pontus: Romeyka in contemporary Trebizond’, Queens’ College, Cambridge, 14 March 2010.
- Chomsky N. (2001). Derivation by Phase. In M. Kenstowicz (ed.) *Ken Hale: a life in language*, Cambridge: MIT Press.
- Bonet E. (1991). Morphology after syntax: Pronominal clitics in Romance. PhD thesis, MIT: Cambridge MA.
- Drettas G. (1997). *Aspects Pontiques*. Paris: Association de recherches pluridisciplinaires.
- Kaltsa M. & Sitaridou I. (2009). The syntax of Pontic dialects. State-of-the-art Survey. Ms. available, University of Cambridge.
- Kaltsa M. & Sitaridou I. (2010). ‘Topicalisation in Pontic Greek’. Paper presented at *Modern Greek Dialects and Linguistic Theory 4*, Chios, June 2009.
- Larson R. (1988). On the double Object Construction. *Linguistic Inquiry* 19: 335-391
- Mackridge P. (1987). Greek-Speaking Moslems of North-East Turkey: Prolegomena to Study of the Ophitic Sub-Dialect of Pontic. *Byzantine and Modern Greek Studies* 11: 115–137.
- Manolessou I. & Beis S.(2006). Syntactic isoglosses in Modern Greek dialects: The case of the indirect object. In A. Ralli, M. Janse & B. Joseph (eds.) *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory*, Mytilene, September 2004, Patras: University of Patras, 220-235.
- McGinnis M. (1999). Evidence for feature-driven A-scrambling. In S. Bird, A. Carnie, J. Haugen, and P. Norquest (eds.) *Proceedings of WCCFL 18*, Somerville, MA: Cascadilla Press, 364-378.
- McIntyre A. (2009). ‘High and low datives in German’. Handout. Invited talk presented at the University of Tübingen. Tübingen, Germany, September 11, 2009.
- Michelioudakis D. (2009). Active vs. Inactive Case features. The typology of dative substitutes in 3 Greek varieties and diachronic implications. Paper presented at ICGL 9, University of Chicago, October 2009.
- Michelioudakis D. (in prep.). Dative constructions in different varieties of Greek. PhD dissertation, University of Cambridge.
- Müller G. (1995). *A-bar Syntax: A Study in Movement Types*. Berlin: Mouton de Gruyter.
- Pylkkänen L. (2002). *Introducing Arguments*. MIT: PhD dissertation
- Sitaridou I. (2009a). Survival against all odds: The last Greek infinitive in Romeyka of Pontus. Submitted.
- Sitaridou I. (2009b). *Language Report on Romeyka*. Ms. available, University of Cambridge.

The impact of Greek on the northern Azov varieties of Russian

ALEXANDER PAVLENKO

GALINA PAVLENKO

Taganrog Institute of Management and Economics,, Russia

alex_pavlenko@inbox.ru

galina21c@inbox.ru

1. Introduction

This paper represents the initial stage of the project being carried out in the field of diachronic sociolinguistics. The project deals with the history of the Greek speech communities of the Northern Sea of Azov coast, especially that of the port-town of Taganrog [Fig.1]. We focus on the outcome of the Greek-Russian language contact in the area which had been developing since the late XVIIIth c. to approximately early XXth c.



Figure 1: *The Sea of Azov*

A peculiarity and a difficulty of the research consists in the fact that there is no Greek speech community in Taganrog any longer. The local Greeks partly immigrated and partly got assimilated having lost their language altogether. That is why the data to be collected is primordially scanty and is being obtained from the censuses carried out before 1917, very scarce traces of the Greek influence in the local Russian speech, if at all, reminiscences of the old-timers, works of belles-lettres and literary memoirs.

The major aim of the project is to reconstruct the development of linguistic situation in Taganrog from the late XVIIIth to the early XXth c. focussing on the interaction between the Greek and Russian communities. It is supposed to compare the social functions of the languages in question as well as to study the nature of the Greek-Russian language contact and its effect on local Russian.

2. Social history

On March 28 1775 Russian Empress Catherine II issued an order acknowledging the Greeks' and Albanians' contribution into the recent victory over the Ottoman Empire. It was also stated that Count Orloff was to organize the settling of Greeks and Albanians in the towns of Kerch and Yenikale in the Crimea as well as in Taganrog on the northern coast of the Sea of Azov [1]. It was not only the considerations of tribute and gratitude to the Greek allies but also the necessity to provide the newly acquired lands with reliable and industrious population that made the Russian authorities to admit the settlers. The settling of these immigrants was by all means desirable for the empire. This incentive also

combined with the long-term plans to re-establish the Greek Empire with its capital in Konstantinopol to be headed by a member of the Russian Royal family.

The Greek and Albanian settlers arrived on board their ships and first started to settle in Kerch and Yenikale as the nearest destination. However the Crimea was not yet under the Russian rule at that time, so the newcomers could not find sufficient land and supplies they needed for a living. Thus, it was decided to settle both Greeks and Albanians homogeneously on the northern Azov coast and to provide them with a considerable amount of money to cover the first needs. Taganrog, which had been re-established in 1769 after being under the Turkish rule and in ruins for 57 years, became the centre of the settlement as the only urban centre in the whole area.

This resettlement started in 1776. The Greeks occupied most lands along the coastline and soon monopolized all the economic activities in the area. That situation was also caused by the fact that the Don Cossacks seized the most fertile lands lying further north of the Sea of Azov thus blocking the way to the sea for the settlers from the inner Russian territories. Therefore Greeks did not have any competitors belonging to other ethnic groups of the area [ibid.].

In the reign of Catherine II Greeks first settled in Kerch and Taganrog and later in Mariupol. However the latter by large became the home for the Crimean Greeks who were mostly the native speakers of the Crimean Tartar language and who acquired the Tartar culture. Only few of the Crimean Greeks were present in Taganrog inhabited by richer settlers of higher social status belonging to military and merchant classes who originated from the Aegean archipelago and continental Greece [See, e.g., Fig. 2]. That majority arrived at Taganrog via Kerch, the poorer of them, mainly fishermen by their occupation, having remained in the Kerch area.



Figure 2: *Ioannis Varvakis (1745-1825) – a Greek national hero, a member of the Filiki Eteria and a distinguished member of the Greek and Russian communities of Taganrog. He spent large amounts of money for construction of Greek Church and Greek Jerusalem Monastery in Taganrog (see below) in the early XIXth c. (unfortunately, both of them were demolished in 1930s).*

The population of Taganrog before 1917, what is quite natural for a port-town, was ethnically and linguistically heterogeneous. Historically the Greek community was an important and authoritative one in economical and social life. The very physical appearance of the town was formed by the architectural tastes cultivated by the Greek population [Fig. 3, 4, 5].



Figure 3: *Greek Church in Taganrog*



Figure 4: *Greek Jerusalem Monastery in Taganrog*



Figure 5: *Alferaki Palace - a mansion built by Nikolay Alferaki (see below) in Taganrog in 1848.*

According to the census of 1872 there were 1807 merchants in Taganrog in those days, among them 334 - Russians and Ukrainians, 481 - Greeks, 242 - Jews, 30 - Germans, etc. The Greek minority gave quite a number of famous tycoons, who made their fortune in and around Taganrog, efficient civil servants, and intellectuals [Fig. 6, 7].



Figure 6: *Alleged portrait of Nikolay (Nikos) Dmitrievich Alferaki, a rich merchant and civil servant (1815-1860).*



Figure 7: Achilles Nikolayevich Alferaki (born in 1846) – Nikolay Alferaki' son; the Mayor of Taganrog from 1880 to 1888.

Judging by the considerable number of local Greeks and the important part they used to play in life of the area it can be taken for granted that there must have appeared numerous Greek-Russian bilinguals in those days. No doubt, this condition was both a result and a motive force of intensive language contact.

3. Language facts

Urban dialects can be seriously affected by either cultural or economical predominance of some social elements over the others. The prestige of a dominating ethnic group's speech may cause imitation on behalf of other ethnic groups. Illustrating this phenomenon A.A.Shakhmatov (1864-1920) mentions some peculiar language developments within the regional varieties of Russian. Thus he mentions that in South Russia, namely in the towns of the Northern coast of the Sea of Azov and in Taganrog in particular, one could come across the examples of transition from affricate [ts] to sibilant [s] in the words like Rus. *tsar'* (czar) pronounced as [sar'] instead of common Russian ['tsar'] [2].

This and other phonetic changes can be attributed to imitation of the Russian speech of the Greek population, which used to dominate the economic life of the area for about a century. In the case of the Northern Azov coast varieties of Russian we deal with an example of mixed urban dialect constituted by the idiom of a quantitative majority (i.e. Russians) on the one hand and that of a minority (i.e. local Greeks) on the other. As is known the local Greeks found it difficult to pronounce the Russian hushing sounds substituting them by sibilants thus producing a "lisp" effect, this feature being sometimes used as a label by the Russian authors (e.g. Anton Chekhov) making the speech of their Greek characters more verisimilar.

From "**The Wedding**" by A. Chekhov (translated into English by Julius West)

CHARACTERS

- EVDOKIM ZAHAROVITCH ZHIGALOV, a retired Civil Servant.
- HARLAMPY SPIRIDONOVITCH DIMBA, a Greek confectioner

The scene is laid in one of the rooms of Andronov's Restaurant

ZHIGALOV [To DIMBA] ...And do you have tigers in Greece?

DIMBA. Yes.

ZHIGALOV And lions?

DIMBA. And lions too. In Russia Zere's nuSSing, and in Greece Zere's everySing – my faZer and uncle and broZeres – and here Zere's nuSSing.

(Russian version)

Жигалов (Дымба). А тигры у вас в Греции есть?

Дымба. Есть.

Жигалов. А львы?

Дымба. И львы есть. Это в России ниЦего нету, а в Греции все есть.

Там у меня и отец, и дядя, и братья, а тут ниЦего нету.

In Dimba's speech in the original Russian version the correct Russian /tʃ/ is substituted by /ts/, while the English translator substitutes /θ/ by /s / and /ð/ by /z/.

The imitation of this "Greek Russian" variety caused some obvious changes both of consonants and of vowels (to a less extent) in the speech of the local Russian population. The most conspicuous changes are as follows:

- 1) some consonants, which are usually hard in other varieties of Russian, in the Russian speech of the Greek community get palatalized as in *pyshka* ['pʲɪʃka] (Eng. *a puff; a bun*) – ['pʲiʃka]; *rynok* ['rɪnək] (Eng. *a marketplace*) – ['rʲinək]; *ryba* (Eng. *fish*) ['rɪbə] – ['rʲibə];
- 2) hushing sounds are substituted by sibilants as in *krysha* (Eng. *roof*) ['krʲɪʃə] – *kryssa* ['krʲɪsə]; *Masha* (diminutive of *Mariya* – a female name) ['mʲaʃə] – *Massa* ['masə], etc.

3.1. Summing up Phonetic Features

According to the available data some peculiarities of the variety of Russian language, which used to be spoken by the Greeks of Taganrog and some other Southern Russian towns are as follows:

- 1) soft post-alveolar affricate /tʃ/ (represented in Russian by letter «ч») turns into hard alveolar affricate /ts/ (represented by letter «ц»);
- 2) hard alveolar affricate /ts/ (represented by letter «ц») turns into hard alveolar voiceless fricative /s/ (represented by letter «с»);
- 3) hard post-alveolar voiced fricative /z/ (represented by letter «ж») turns into hard alveolar voiced fricative /ʒ/ (represented by letter «з»);
- 4) hard post-alveolar voiceless fricative /ʃ/ (represented by letter «ш») turns into hard alveolar voiceless fricative /s/ (represented by letter «с»);
- 5) Russian close central vowel /ɨ/, which indicates hardness of the previous consonant is substituted by close front /i/ indicating palatalization of the previous vowel;
- 6) Palatalization might occur also in certain contexts.

4. Conclusion

The collected evidence shows that at one time South Russian town-dwellers, especially women, might have started to imitate the abovementioned phonetic features of the local Greeks' speech considering it prestigious. These features spread in the varieties of Russian all along the Northern Azov coast as well as in the Cossack towns of the Lower Don area [3]. Quite soon the features in question turned into characteristic peculiarities of the local accents, i.e. they were not mere imitation any longer.

In the late XIX c. some Russian scholars considered the said features to be the direct heritage of the local language contact during the Greek colonisation in the ancient times. However, this hypothesis seems to be hardly probable as language continuity in the Southern Russian steppe region had been broken intermittently because of massive migrations and long periods of devastation and abandonment.

In future it is supposed to carry on a retrospective study of age, sex, and occupational variation as regards the features of Russian attributed to the Greek influence. An approximate estimation of the time by which this accent had formed would also be quite tempting. These goals are quite a challenge as most peculiarities under consideration are extinct by now.

References

1. Phielevsky P. P. (1898). *A history of Taganrog*. Moscow.
2. Shakhmatov A. A. (1916). *An Introduction to the course of history of the Russian language*. Part I. Petrograd.
3. Savelyev E. P. (1908). *The Types of the Don Cossacks and the Peculiarities of Their Speech*. Novocherkassk.

MBU!

On *wh*-objects and true adjuncts of Cypriot Greek

NATALIA PAVLOU
University of Cyprus
nataliapavlou@gmail.com

1. Introduction¹: *embu* and *mbu*

This paper starts out from the discussion of the different approaches to the formation of *wh*-questions in Cypriot Greek which involve the use of *embu* and the possible assumptions that have been made for the analysis of *mbu*, an element that may appear having as a host the *wh*-phrase *inda*. It explores the observation that the dialectal *wh*-phrase *inda (mbu)* can have four possible allomorphs which appear to be the result of language change and therefore, present their own morphosyntactic properties which differ from the aforementioned *inda mbu*. The possibility of language change in these *wh*-phrases has been the immediate observation of a questionnaire, examining the syntactic restrictions among the allomorphs in four different age groups. The final section of this paper proceeds to show how these four allomorphs are different from the standard form by taking into account any phonological and morphosyntactic properties and by exploring different syntactic analyses for the standard form and its apparent allomorphs.

Embu and *mbu* are some of the most obvious markers for Cypriot Greek and therefore, have been extensively used in texts which are included in books discussing the Cypriot Greek history (Simeonidis 2006). The optionality in forming *wh*-questions in Cypriot Greek by using *embu* or not has been a significant matter of recent discussion in the literature of Cypriot Greek. (Grohmann, Panagiotidis and Tsiplakou 2006, Papadopoulou in progress). Cypriot-Greek speakers have the optionality of using an extra element *embu* in *wh*-questions introduced with *wh*-arguments (both subjects and objects), *wh*-quasi-arguments and true adjuncts:

- (1) a. Pcos embu emilisen?
Who embu talked.3SG
'Who talked?'
- b. Pcos emilisen?
Who talked.2SG
'Who talked?'

Grohmann, Panagiotidis and Tsiplakou (2006) suggest an analysis assuming sideward movement in a cleft structure whereas Papadopoulou (in progress) argues that *embu* is a fossilized element meaning that its past structure might have been a more complex one but it has been simplified in one element through the passing of the time and can only appear in the Complementizer position.

This paper deals with *mbu*, a variant of *embu* which appears in different contexts obligatorily and may support different functions. The relevant discussion for this paper involves the obligatory use of *mbu* in *wh*-questions, where *embu* is not allowed. One of the most important differences between the two was observed by Grohmann, Panagiotidis and Tsiplakou (2006) in complex *wh*-expressions with *inda* and a noun phrase, where

¹ I express my gratitude and admiration to Kleanthes Grohmann, with whom this topic originated as my linguistics research paper, for his continuous encouragement and the support that he always offers to students of all levels as well as his endless discussions and assistance I enjoyed myself, which also helped me identify the properties of the *mbu*-allomorphs and provide further explanations.

there is obligatory use of *embu* (2a) and its contrastive use when *inda* is used as an argument and it necessarily needs *mbu* (2b). This observation holds for both main and embedded clauses:

- (2) a. *Inda fain {embu, *mbu} emairepses?*
 What food.ACC *embu* cooked.2SG
 ‘What food did you cook?’
 b. *Inda {*embu, mbu} emairepses?*
 What *mbu* cooked.2SG
 ‘What did you cook?’
- (3) a. *Pe mu inda fain {embu, *mbu} emairepses*
 Tell.2SG me.ACC what food.ACC *embu* cooked.2SG
 ‘Tell me what you have cooked’
 b. *Pe mu inda {*embu, mbu} emairepses*
 Tell.2SG me.ACC what *mbu* cooked.2SG
 ‘Tell me what you have cooked’

A second difference between the two, which can be argued to play a role for the claims of this paper, is the exceptions to the *embu*-strategy. The Standard Greek *wh*-phrases *ti* ‘what’ and *jati* ‘why’ cannot be combined with *embu* but, as it appears, *mbu* and its host *inda* are used as the only alternative option to the ungrammaticality noted below:

- (4) a. **Ti embu efaes?*
 What *embu* ate.2SG
 ‘What did you eat?’
 b. *(?)²Jati embu epies?*
 Why *embu* went.2SG
 ‘Why did you go?’

The structure in (4a) is unacceptable and its grammatical form would appear with *inda mbu* (5a), where as the structure in (4b) is considered ungrammatical by a significant number of Cypriot speakers, who have claimed that (5b) would be a more preferable way of forming the question:

- (5) a. *Inda mbu efaes?*
 What *mbu* ate.2SG
 ‘What did you eat?’
 b. *Inda (mbu) epies?*
 Why (mbu) went.2SG
 ‘Why did you go?’

This can be a matter of combining Standard Modern Greek *wh*-phrases with a purely Cypriot-Greek element resulting in a mixing of the two.³ This appears not to simply be

² The single question mark indicates mild ungrammaticality or grammaticality by a specific set of people.

³ See also Fotiou (2009) for a relevant discussion on the ungrammaticality of the combination of Standard Modern Greek (SMG) and Cypriot Greek (CG) regarding structural focus and Panagiotidis (2009) for relevant comments on the morphological and syntactic mixing in CG. For relatively opposite effects, there is recent work on clitics by CAT (Grohmann, Theodorou, Pavlou & Leivada 2010), a recently-founded research team (Grohmann 2009), which concentrates on the mixing of SMG and CG due to external factors and the implications on the structure of CG.

code-switching, but the use of both elements between Greek and Cypriot Greek which results to unnaturalness of the sentence. In this paper, it will be argued that in these cases, there is use of *inda* and its follower *mbu* along with the appearance of the allomorphs among the younger population. By this, it is implied that the use of the latter is much more frequent than the use of the Modern Greek *wh*-phrases *ti* “what” and *jati* “why” combined with any Cypriot-Greek expressions and as it will be shown later on, this has given a new shape to Cypriot-Greek *wh*-questions.

A third difference is related to *wh*-questions where *mbu* along with its host *inda* seem to attract other elements, a property also found in *embu*-questions. The following examples show that *mbu* in copular sentences attracts the Cypriot copula *en/eni*:

- (6) a. Pcos emboni?
 Who embu is.3SG
 “Who is it?”
 b. Pcos embon tzinos?
 Who embu en.3SG he.NOM
 “Who is embu he?”
 c. Inda mbon / Inda mboni?
 Inda mbu en.3SG/ Inda mbu eni.3SG
 “What is it”
 d. Inda mbon tzino?
 What mbu en.3SG it.NOM
 “What is that?”

Supposing that verbs raise at least to T⁰ in Greek and possibly in Cypriot Greek as well, then the copula lands in T⁰ as well. Following Papadopoulou (in progress) that *embu*, and logically its variant *mbu*, are Complementizers, it can be assumed that the kind of close distance between the copula in T⁰ and (*e*)*mbu* in C⁰ has the phonological effects of *mboni/mbon* (*mbu+ eni/mbu+ en*).

2. Exploring the *inda/ inna /na/ ta/ a mbus*

2.1. The *inda mbu*

Even though *embu* and *mbu* show some similarities in their structure, the fact that they appear in different structures cannot be ignored. This section will be discussing the properties of *inda mbu* ‘what’ and ‘why’ and present some of the tests and restrictions that explain the special nature of *mbu*.

The close relation of ‘what’ and ‘why’ is not surprising, since *ti* ‘what’ can take the role of *jati* ‘why’, as shown below:

- (7) Ti to ekruses?
 What. it.ACC burnt.2SG?
 ‘Why did you burn it?’

This kind of constructions is very often in CG- and respectively, in other varieties as well. Even though the two are syntactically very different, they appear to share a lot of similarities in the proposed topic. ‘Why’ appears to show similarities with ‘how come’, as Tsai (2008) explains for *why-how come alternations*, which although on a first glance seem of the same nature, they show a lot of syntactic differences and dependencies.

A first look at *mbu* was first introduced by Grohmann, Panagiotidis and Tsipplakou (2006) who observe that *mbu* is used obligatorily when serving with *inda* having the function of an argument.

- (8) Inda mbu vastas tziame?
 what *mbu* hold.2SG there

'What are you holding there?'

Inda is believed to have originated from the interrogative pronoun *tinda*, used in Asizes (Simeonidis 2006, mentioned in Grohmann & Papadopoulou to appear). As far as its today's use is concerned, it appears that some minorities in certain regions of Cyprus which show more dialectal heaviness than other areas use the *inda* 'what', where as most of the population today does not, suggesting possible language change.

- (9) *Inda mairefkis?*
 What cooking.2SG
 'What are you doing?'

Inda 'what' in those minorities shows some interesting structures, which are not shared by the rest of the population:

- (10) *To master sta linguistics inda na to kamo?*
 The master.NOM in linguistics what to it.ACC do.2SG
 'What would I do a master degree in linguistics?'

In (10) there is *wh*-movement out of a predication relation, already identified as a possibility in SMG (Spyropoulos 1999), meaning that the answer to this question would be (*kame to*) *kadro* '(do it) a picture'. Contrary to this, the *inda* in this kind of structure would be an adjunct for most of the Cypriot speakers today.

Other than this, *inda* 'what' is widely used in "frozen expressions", indicating the possibility of language change and loss of it in today's language, and its remaining through cultural specificities expressed by these expressions:

- (11)a. *Inda kori?*
 What girl.NOM
 'What's up girl?'
 b. *Inda kamnis?*
 What do.2SG
 'How are you?'

Even more interestingly, this kind of expressions can also be found with *na*-clauses and certain verbs in cases which may fairly be called 'echo-questions' in populations, where *inda* 'what' is not grammatical⁴:

- (12)a. *Inda na kamo?*
 Whatna do
 'Do I have another choice'
 b. *Inda na su kamo?*
 What na you.GEN kamo.1SG
 'I can't do anything for you'
 c. *Inda na pis?*
 What na tell.2SG
 'There's nothing to say!'

But, not:

- (13)* *Inda na su goraso?*

⁴ In the minorities where *inda* 'what' is grammatical, sentences in (12) can also have the literal meaning.

What na you.GEN buy.1SG
 'What do I buy for you?'

These fixed meanings, in a non-idiomatic way, that the echo-questions have and the 'survival' of *inda* 'what' in minorities is assumed here to be the support for its change, or even death.

Mbu shows optionality even today, when combined with *inda* serving as an adjunct:

(14) *Inda (mbu) me thoris?*
 why *mbu* me.ACC look.2SG
 'Why are you looking at me?'

Two tests, the negation and the DP-test, have been applied to identify differences between the 'why' and 'what' or the bare form without the *mbu*:

(15)a. *Inda en efaes?*
 Why not.NEG eat.2SG
 'Why did you no eat?'

b. *Inda mbu en thelis*
 What/Which *mbu* not.NEG want.2SG
 'What do you not want'

c. (?) *Inda mbu en efaes?*
 Why *mbu* not.NEG eat.2SG
 'Why did you not eat?'

As can be seen in (15c), the mild grammaticality of the negation⁵ with the adjunct *wh*-phrase comes in oppose with the perfectly correct questions with the *wh*-object in (15b). This already suggests that there can be some differences between the two. Agouraki (2010) discusses the emphatic role of Neg-to-C as an element expressing an [Emphasis] specification on the fill-requirement of C. If *mbu* is a variant of another complementizer (Papadopoulou in progress) as discussed in the first section of this paper, then the already taken position by the negation in C causes the derivation to crash. However, since this is only one example, I will not argue at this paper for the structural position of negation in CG. As striking as it may seems, the *wh*-object *inda mbu* brings no objections to negation revealing that there are indeed some difference between *wh*-object and true adjunct, which will be discussed later on.

Another test that was put in use to expand the already existed knowledge and reveal the nature of *inda mbu* was the DP-test, as will be called here, where the determiner takes the position of the D head and gives the following:

(16)a. *To inda mu eklepses ta lefta en ekatalava.*
 The why me.GEN stole.2SG the money.ACC not.NEG understood.1SG
 'The why you stole my money I did not understood'

b. (?) *To inda mbu mu eklepses en mu ipes*
 The what *mbu* me.GEN stole.2SG not.NEG me.GEN said.1SG
 'The what you stole from me you haven't told me'

c. (?) *To inda mbu mu eklepses ta lefta en ekatalava*
 The why *mbu* me.GEN stole.2SG the money.ACC not.NEG understood.1SG
 'The why you stole my money I did not understood'

⁵ I thank Anastasia Giannakidou for sharing her thoughts with me on this issue and Anna Roussou for pointing negation as a possible test for clarifying the *mbu*-allomorphs.

Wh-phrases have the property of becoming determiner phrases (DP) (Abney 1987) when a determiner is placed in D. While all the rest of the *wh*-phrases in Cypriot Greek (i.e. *pcos* ‘who’, *pote* ‘when’, *pou* ‘where’, *jati* ‘why’, *ti* ‘what’ etc.) and *inda* ‘why’ share this property, the *inda mbu* (both object and adjunct) are accepted by some speakers or even by those accepted they do not sound very grammatical. The observations here may not result simply from the existence of a Complementizer but, from the combination of *inda* and *mbu*, with *inda* being a fused form resulting to a cleft (with *mbu*), since its literal meaning is *ine ti afta* (Pavlou in progress, Grohmann and Papadopoulou to appear).

- (17)a.* To ine ti (inda) pu efaes den mu ipes
 The is what that ate.2SG not me.GEN told.2SG
 ‘The what you ate, you didn’t tell me’
 b. To ti en pu (embu) efaes, den mu ipes
 The what is that ate.2SG, not me.GEN told.2SG
 ‘The what you ate, you didn’t tell me’

2.2. The *mbu*-allomorphs

Interestingly enough, *innambu*, *nambu*, *tambu* and *ambu* which are claimed here to be the four possible allomorphs of *mbu* do not share the same morphological properties as the *inda mbu*, which will be called here the standard form of use on the island. A closer look at them reveals that the phonological similarities with *inda mbu* are only at a first glance but, this is not the only case as illustrated below:

- (18)a. To moro {innambu, *inna} klei?
 The baby why cries.3SG
 ‘Why is the baby crying?’
 b. {Nambu, *Na} fonazis?
 Why shout.2SG
 ‘Why are you shouting?’
 c. {Tambu, *Ta} ekatharises to trapezi?
 Why clean.2SG the table
 ‘Why did you clean the table?’
 d. {Ambu, *A} skupizis to patwma?
 Why sweep.2SG the floor
 ‘Why are you sweeping the floor?’

As observed above, *mbu* is attached to the allomorphs not only when they are used as *wh*-arguments but also as *wh*-adjuncts, resulting to their status as one word. On the contrary to *inda mbu*, the *mbu*-allomorphs cannot be separated in two words and therefore *inda* is no longer considered a host and *mbu* its attached element in *wh*-questions, but the two of them inseparable pieces of the actual *wh*-phrase. So, the allomorphs are lexical items used in *wh*-questions, both *wh*-arguments (objects) and true adjuncts.

This would explain the ungrammaticality of (4) with *wh*-phrases *ti* “what” and *jati* “why” which cannot be combined with *embu* and the existence of the *mbu*-allomorphs or the standard form in their position. If all of them can function as *wh*-objects or adjuncts, then the immediate question regarding *innambu/ nambu/ tambu/ ambu* would be whether there are any syntactic environments where any of these can behave as arguments or adjuncts and if there can be any other disambiguation point, except the meaning of the context. The obvious response would be that their function is determined from the verb’s transitivity determining the variant’s function as arguments or not. In (19a) the clitic in genitive *tu* leads to the immediate identification of *innambu/nambu/tambu/ambu* as the adjunct and in a similar way in (19b) the allomorphs have the meaning of “what for”. In (19c) the allomorphs are used as *wh*-arguments, whether that means landing in Spec, CP or somewhere else. The transitivity which

determines the actual function of the variant being an argument is the first point of disambiguation of the allomorphs:

- (19)a. {Innambu, Nambu, Tambu, Ambu} tu fonazis?
 Why him.GEN shout.2SG
 ‘Why are you shouting to him?’
 b. {Innambu, Nambu, Tambu, Ambu} ton thelis
 What for him.ACC want.2SG
 ‘What do you want him for?’
 c. {Innambu, Nambu, Tambu, Ambu} thelis?
 What want.2SG
 ‘What do you want?’

However, the ambiguity becomes obvious in a sentence like the following:

- (20)a. Innambu/Nambu/Tambu/Ambu fonazusin?
 What/Why shout.3PL
 #1 ‘What are they shouting?’
 #2 ‘Why are they shouting?’

The verb in (11) can be listed as an optionally transitive verb in Cypriot Greek and result in the ambiguity of the allomorphs meaning ‘why’ or ‘what’. If the question was formed with the standard form, namely *inda mbu* then it would most probably be interpreted as an argument (although it can also function as an adjunct) since the most common question that would be asked for the *wh*-phrase to be interpreted as ‘why’ would be *inda fonazusin*. *Inda* shows more frequency of use in Cypriot Greek and this, as will be shown later on, seems to be a determining factor for the allomorphs as well.

Regarding the other properties of *inda mbu* mentioned above, it should be noted that, although *innambu/nambu/tambu/ambu* can function as *wh*-adjuncts and be similar to *inda* or serve as *wh*-arguments meaning ‘what’, they cannot be combined with a complex *wh*-phrase of the type *inda* +noun, as in (2a, 3a). This results that the variants cannot serve as referential *wh*-phrases after their fusion with *mbu*:

- (21)* Nambu fai emairepses?
 What food cooked.2SG
 ‘What food did you cook?’

However, the *mbu*-allomorphs share similar properties to the standard form, like those mentioned in (6c,d), showing that *mbu* is the strongest element between *inda* and *mbu* but still having the unity of the allomorphs as their main property:

- (22)a. Nambon/ Innambon/Tambon/Ambon?
 What is.3SG
 ‘What is it?’

Moreover, they seem to follow the same pattern in the negation test and show the same oddness with the *nambu*-adjunct. Regarding the DP-test, the same effects are also present.

The different properties of *mbu* discussed here show some basic similarities and differences between *embu* and *mbu*, but create the question of ambiguity in the allomorphs. The description of the study following below, aims to unfold any restrictions related to the *mbu*-allomorphs, specify their exact environment and lead to a clearer picture of the *mbu* jungle.

3. The study

Before giving the description of the *mbu*-allomorphs study, it should be pointed out that Cypriot Greek does not have a written alphabet, but rather if there is any in poems, text messages or any other informal form of communication, it is the individual transcription of its sounds using the Greek alphabet and therefore can vary in many levels. The data given for judgment in written form were crosschecked for their naturalness with several speakers before the distribution of the questionnaire who agreed upon some of the sounds which are specifically used in Cypriot Greek.⁶ A sample of these is given in (13) while the rest of the sounds follow basic transcription of Greek in general:

- (23) /ts/ i.e. τήνος ‘that one’
 /sh/ i.e. έση ‘(it) has’

The statement above, also mentioned in many works on Cypriot Greek (among others Fotiou 2009)⁷ can be listed as a problematic aspect of this study since participants were asked to judge not only the grammaticality of a syntactic order, but the written system itself. However, the majority of the words was spelled following the spelling judgments from speakers and therefore, did not create any serious problems throughout the whole process.

To collect clear competence data is one of the most difficult tasks that a study has to solve and fairly enough there has been strong criticism for the use of questionnaire in doing so. The main concern of a questionnaire is to actually make the participants judge the sentence in front of them, like they would have produced it and not what should be the correct form. The same effort was made for *mbu*-allomorphs following a methodology⁸ with the use of a pen-and-pencil questionnaire to elicit judgments from 100 native speakers, all of them non-linguistically trained. The questionnaire involved both 41 closed test sentences and 10 fillers in order to counterbalance habituation effects like the easiness in informants’ judgments when they get used to a given construction that is being repeated. The small number of fillers can be argued to be the second main weakness of this questionnaire, although there has been no problem observed for the participants in this questionnaire and the number of the constructions tested allow for a small number. Test constructions were randomly put in order and the choice of words aimed to the most dialectal form of them and therefore there was limited use of common words between Cypriot Greek and Standard Greek. Generally in variation studies, texts should be as closer as they can to normal speech and even use vernacular forms (Montgomery 1997). The participants had to choose between a 5-grade scale ranging from *completely unacceptable*, *below satisfactory*, *satisfactory*, *quite good* and *absolutely satisfactory*. The choice of the 5-scale was made on the basis that the 3-scale may not provide the different levels where a sentence can be appropriate, especially within different contexts. In this case, it appeared to be the case that the 5-grade scale was used to judge attitudes of the participants. Grammaticality is more empirically adequate and valid when it is presented in many levels and not binary and for this reason a simpler scale of grammatical/ungrammatical was ruled out but at the same time any larger scale above 5 would be confusing. The

⁶ For a different type of encoding Cypriot Greek sounds see Simeonidis (2006: 375).

⁷ Fotiou (2009) gives a detailed description of the status of Cypriot-Greek, where she mentions specifically the linguistic nature of Cypriot Greek as a dialect, or second variety spoken in Cyprus. Also, Grohmann & Papadopoulou (to appear) briefly discuss the Cypriot context and Ioannidou & Pavlou (2009) present the poverty in Cypriot population’s perception and judgment for their variety.

⁸ Here, I would like to thank Elena Papadopoulou for her willingness to guide me properly through methodological issues and weaknesses of a questionnaire-based study.

participants were selected from the region of Limassol to restrict any regional variation, something which appeared to have significant results for the *mbu*-allomorphs.

There were four syntactic environments being tested which involved clause-initial position of the *mbu*-allomorphs, initially assuming that this is in Spec, CP, topicalized elements i.e. noun phrases, adjective phrases and adverb phrases preceding the *mbu*-allomorphs and last, the *mbu*-allomorphs in embedded contexts and in both declarative and interrogative sentences. The targeted responses aimed to show that there is difference in the syntactic distribution between the *mbu*-allomorphs and also with *inda mbu* which could be related to their morphological difference with it.

4. Setting off

A pilot study administered to 10 adults from Limassol using the same questionnaire as described above gave enough evidence to claim that *innambu* is used with a topicalized element rather than in the clause-initial position. *Nambu* appeared with preference in the clause-initial position, where as the other two, *tambu*⁹ and *ambu*¹⁰, appeared not to be used in the region of Limassol. The distinction that the data of the pilot study draw for the syntactic differences between *innambu* and *nambu*, at least, were a good start to go on with bigger corpus.

Although the pilot study excluded *ambu* and *tambu* as allomorphs used in Limassol, they were not excluded from the questionnaire later on. However, for the purposes of this paper there will be mainly focus on *innambu* and *nambu* which were analyzed from the corpus collected. The full study with the 100 native Cypriot speakers showed the following for each of the allomorphs:

List of General Results:

Innambu

Innambu showed a strong preference by two age groups in its use with a topicalized element either a noun phrase or an adjective phrase or both.

Nambu

In contrast, to the findings of the pilot study, the full study showed no important distinction for the syntactic distribution of *nambu* but, instead participants find it grammatical in any of the environments tested, with a slightly increased preference in clause-initial position.

Tambu and Ambu

Tambu and *ambu* showed low use in comparison with the first two.

Based on the fact that two out of four allomorphs showed some evidence for the targeted responses that the variation and the inconsistency in the data concerning the two cannot be simply the result of inadequate empirical methods, but evidence for regional variation, as mentioned above. Interestingly, although the observations above point to an important distinction between *innambu* and *nambu*, these were only true when they were used as *wh*-arguments. When either one of the two was used as adjunct, then there was no difference in the syntactic environments noted by the participants. This leaves implications for the *wh*-arguments and *wh*-adjuncts in Cypriot Greek, which will be discussed later on.

4.1. Attitudes for language change

One of the most significant findings of this study is the sociolinguistic status of the *mbu*-allomorphs which was shown by the age factor of the participants. As mentioned

⁹ As informed by participants *tambu* is used in rural regions.

¹⁰ *Ambu* was very strongly claimed by a big number of participants that it is widely used in the region of Paphos, the southwest part of Cyprus.

above, the participants were grouped in ages of 18-30, 30-45, 45-60 and 60+. Based on these ages, the results imply that there are attitudes for change, starting from no use at all of *nambu* and gradually increasing till the age of 18-30, where there is use of *nambu*:

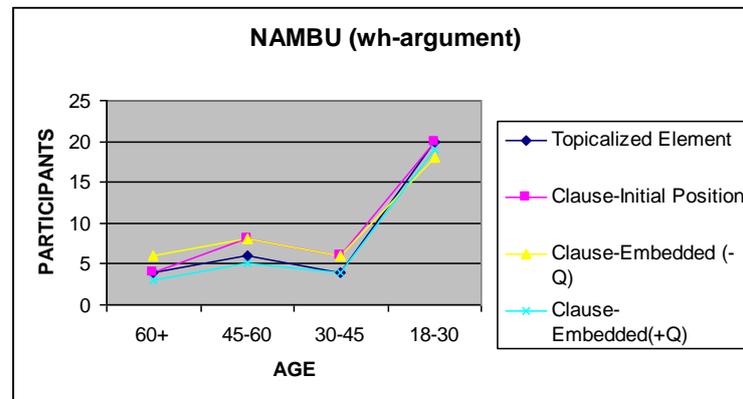


Figure 1.

There is a slight increase at age 45-60, which falls again at the age group of 30-45 and then rises to give the 80% of the test sentences given as grammatical with *nambu* in all the environments tested. Possibly, the age of 30-45 shows a fall on the use of the allomorphs since this is the age of parents raising children and in the Cypriot context, this implies that they would speak 'proper' Standard Modern Greek to the children.

The data provided for *nambu* shows immediately the observation of ongoing language change. Papadopoulou (pc) also notes appearance of the *mbu*-allomorphs in younger children (of age 2;0-3;0) in spontaneous speech. Since Labov's success of his methodological innovations in Martha's Vineyard (1963) and in New York City (1966), linguistic research has been following Hockett's (1958) confirmation that the actual process of language change can only be detected through the result of this kind of studies. Over the last 30 years, language change can be analyzed during the period that is happening. The apparent-time construct which can be characterized as the quickest, easiest and safest way of replacing real-time data has been one of these important Labovian innovations, which can take into account the linguistic variation that appears before language change.

Bailey (2002) reports that age is statistically significant for each variable but it cannot always predict that there is ongoing language change and not "stable variation". Change follows prototypically a path where some variant in the speech of older group in the community appears more frequent in the speech of the middle generation and even more in the youngest generation. Although figure 1 does not look like the characteristic shape of S-curve graphic representations that are known for language change (Weng and Cheng 1970, Chambers & Trudgill 1998), the claim is that the three stages of language change- initial stasis, rapid rise and tailing off are not all captured through this sample. The figure represents an idiosyncratic way of language change, in the sense that there is long and almost steady initial stasis in the ages 30-60+ and a very late finishing with a sudden acceleration of the young group. The rapid rise does not appear at all or if it does, it can only be characterized as sudden, since there is no steady rise for the descending ages of the subjects participating. This can mean a) that the data collected capture the first stage of language change showed by the initial stasis of the 30-60+ or b) if this is the first stage of change, the sudden acceleration of the youngest group is only a rise of frequency of variation which has been argued to occur before language change so as the new elements attain some kind of critical mass (Chambers 2002). The problem is that this rise in frequency has been reported to be gradual and really difficult to observe but taking into account that 20 young people reported above the use of this variant is very much clear to all. Charts of similar type have been shown for the Dialect Topography of Canada

(Chambers 1994), where Quebec City appears to take an idiosyncratic path in the middle part of the change. After the initial stasis, there is sudden acceleration to change in the 40-years old participants following the kind of pattern noted in Figure 1. For sure, if this is a change taking place for Cypriot Greek, it is progressing very rapidly; and this does not characterize a well-behaved language change.

It is worth noting that the small number of data collected for *tambu* show a normal increase in the use of it in the speech of younger people. It presents a steady rise for the age groups, moving from old to young, capturing exactly the apparent-time construct effect. *Ambu* shows many idiosyncrasies in the different structures tested. As far as topicalized elements are concerned it presents similar sudden acceleration with *innambu*. For the clause-initial position, *ambu* behaves normally and the change happens gradually giving the S-shape. For the embedded contexts, the initial stasis seems to hold for the age 40-60+ and then the language change starts in normal pace. The charts are not given because the numbers of the data collected are not representative, since the two allomorphs are not used in Limassol Cypriot Greek or if they are used the corpus collected is not adequate to account for any generalizations. The observations mentioned above for *tambu* and *ambu* can be taken as tendencies or behaviors, which are the only safe observation that can be taken out of the two.

The case of *innambu* brings another issue into discussion. It would be the same with *nambu*, if there wasn't this abnormal use of *innambu* with a topicalized element in the age group of 45-60, which declines and then rises again. The use of *innambu* with a topicalized element was the targeted construction from the start and although it was captured, it shows some strange patterns which are described below. The pattern in Figure 4 creates the question of age-grading and whether this particular construction is repeated in different phases of life. Since this paper follows the hypothesis of the apparent-time construct which does not include age-grading, there is no obvious reason that Cypriot Greek speakers alter their way of speech to adopt some norms in the age of 45-60 and 18-30. For age-grading to be argued, there must be even clearer data.

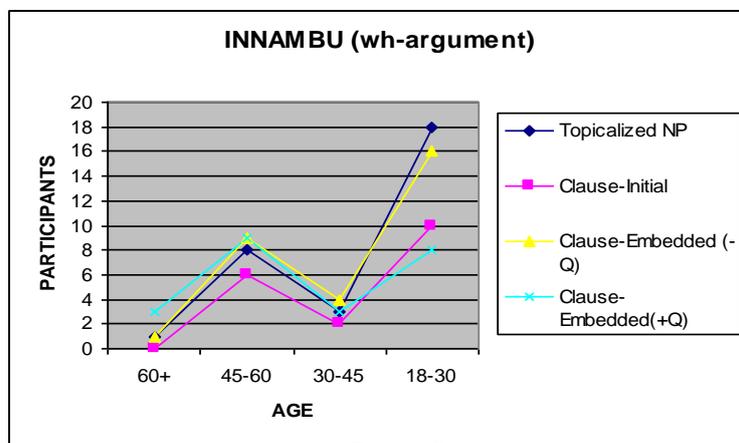


Figure 2.

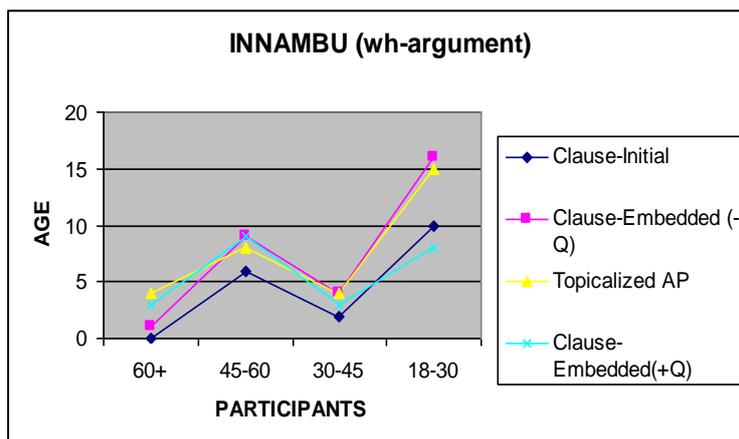


Figure 3.

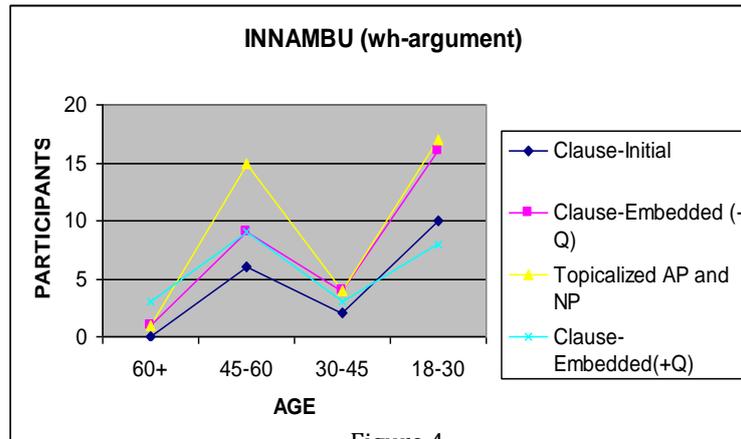


Figure 4.

A second prediction would like the pattern showed below to remind Labov's study in Martha's Vineyard, where two age groups had roughly similar scores, and the other groups having very different scores. Well-known by now is the similarity of less frequency in the use of the variable tested between the 61-75 and 14-30 age groups on the island of Martha's Vineyard. In the same way, *innambu* which is argued in this paper to be another variant under language change shows increased frequency of use in the age groups of 45-60 and 18-30. The charts in Figure 2-4 present the different topics given for *innambu* as the targeted structure (NPs, APs, or both) and how all three follow the same patterns.

Comparing these charts to *nambu* in Figure 1, the conclusions are very much different. There is no stasis at all, as shown for *nambu* and the increase in frequency and use is not only observed in the youngest group but in a strange way in two groups. Whatever the social reasons for the similarities between 45-60 and 18-30 are, they are of no special importance to this paper, but there is one clear point to be made: The "reversed Vs" in the charts for *innambu* show that the variant is used in different ages. If this is not to be taken as age grading and logically loss of the variant at some point, then by concentrating in the youngest age group, there can be a tendency for language change.

Whatever the reasons are the apparent-time differences noted among generations of the Limassol Cypriot Greek mirror diachronic developments in language and imply some attitudes towards change going on in 'real time'. Studying language change diachronically is for sure the ideal method (Labov 1982) but, it can only happen when someone re-interviews the same individuals over a period of years. The methodology of the questionnaire used here rules out this possibility because of its anonymity so the best assumptions can be made by looking into this corpus collected.

5. Variation and Syntactic Theory

The question relating *mbu*-variation and syntax is yet to be answered. The *mbu*-allomorphs show a status that does not involve being determined by any social factors, rather than just being element that are currently changing. Indeed, regional variation (Limassol, Paphos etc.), especially for *tambu* and *ambu*, can be argued to be related to a particular group of people, but still this can leave no implication for stylistic aspects or external factors, since regional variation cannot be seen as style dependent. So, any assumption that can be made for the *mbus* as phonological allomorphs based on the speaker's performance can be ruled out at this point. However, there is one question remained to be answered: Should the difference in frequency of use of these allomorphs account for variation in syntax?

The data show that individuals make use of allomorphs varying in frequency and frequency is very logically related to everyone's mind with stylistic aspects. Henry (2002) suggests that variation can be syntax's job and as these data show variation is not necessarily linked to any stylistic factors. What is clear is that if any assumption of language change can be taken into account, then the issue of frequency is the first to be considered. Based on the ages of 18-30, young speakers of Cypriot Greek have just started making use of these allomorphs but, at the same time have in their grammar the standard form then there should be expected a decline of it and more use of the allomorphs. As shown in Figure 5, there is slight fall of the use of *inda mbu*, which can only show a tendency and cannot be considered as evidence:

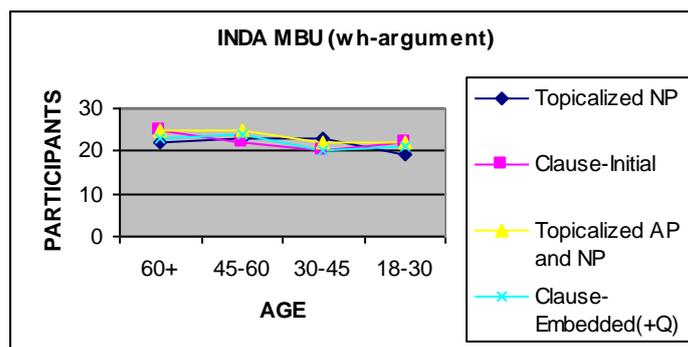


Figure 5.

However, the graphs given so far show use of *inda mbu* and *nambu* in the group age 18-30, as well as *innambu* with topicalized elements and as Complementizer introducing embedded clauses, and *ambu* and *tambu* in much less percentage but still in use by the same age group. This can only show variation in the grammar, which is not marked by any stylistic factors but a rare regional distribution on the island, which can be doubted, and differ in frequency of use. Whatever the case is for the two (or four) allomorphs, the picture created from this corpus is that these two may be under regional variation, something that will be confirmed once a similar study is carried out in other parts of Cyprus. Judging from oral data, it seems that the two allomorphs are not only used in the region of Limassol. If this happens, then these allomorphs have a status of free variation in syntax. Assuming that external factors (distraction while filling the questionnaire, Cypriot Greek lacks a written alphabet etc.) did not play any role to have these results, and the *mbu*-allomorphs are to be listed as part of the *competence* then a first problem comes down to the issue of a grammar allowing different frequencies for each of these allomorphs, as already mentioned above.

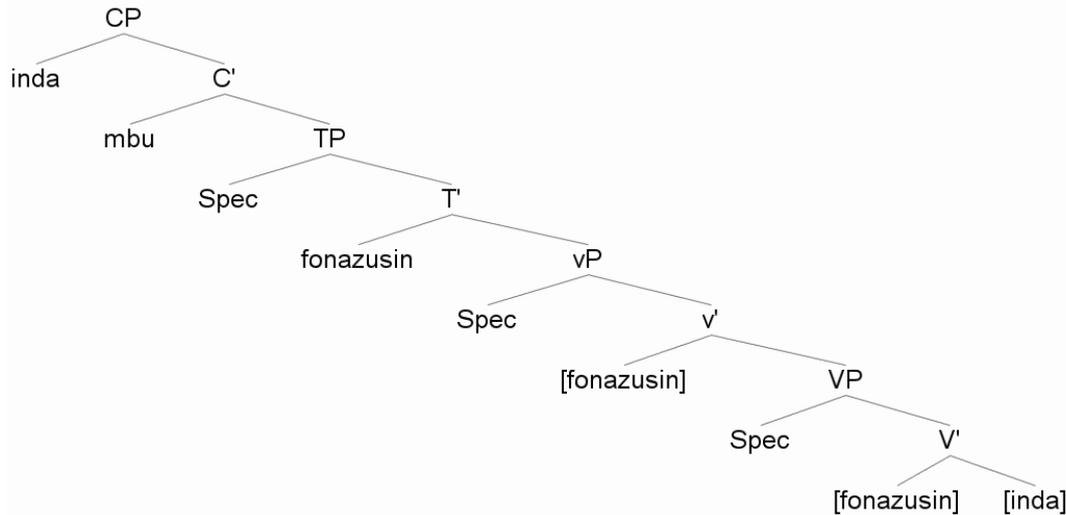
For sure, what can be excluded for the moment is that the *mbu*-allomorphs are not elements of an idiolect because the choice is not personal based on different social factors. The data collected show that a person can use both the standard form *inda mbu* and the allomorphs *nambu* and *innambu*, without any importance to register at all. So, what can be assumed is that these allomorphs used interchangeably for the time being is an immediate result of the co-existence of all of them in grammar. If language change is indeed taking place, then there should be expected to see in future work more syntactic restriction, like the case of *innambu*.

5.1. The *mbu*-structure

Having clarified that the *mbu*-allomorphs are new elements in Cypriot grammar, there should be a syntactic representation which illustrates the different scenarios of the *mbu* puzzle. Before moving into the structure of the allomorphs, it is necessary to discuss the structure of the standard form of *inda mbu*, for the sentence given in (20), repeated here as (24). Even though there is not any relevant work on the structure of *inda*, there are possibilities easily observed to any Cypriot which would suggest *inda (mbu)* being a fused

form of a cleft *ti ine (pu)* ‘what is (that)’ or *ine ti pu* ‘is that what’. Under the hope of a future study investigating this (Pavlou in progress), *inda* will be used in Spec, CP for the purposes of this paper since the concentration lies on *mbu*.

- (24) *Inda mbu fonazusin*
 What *mbu* shout.2PL
 ‘What are they shouting?’



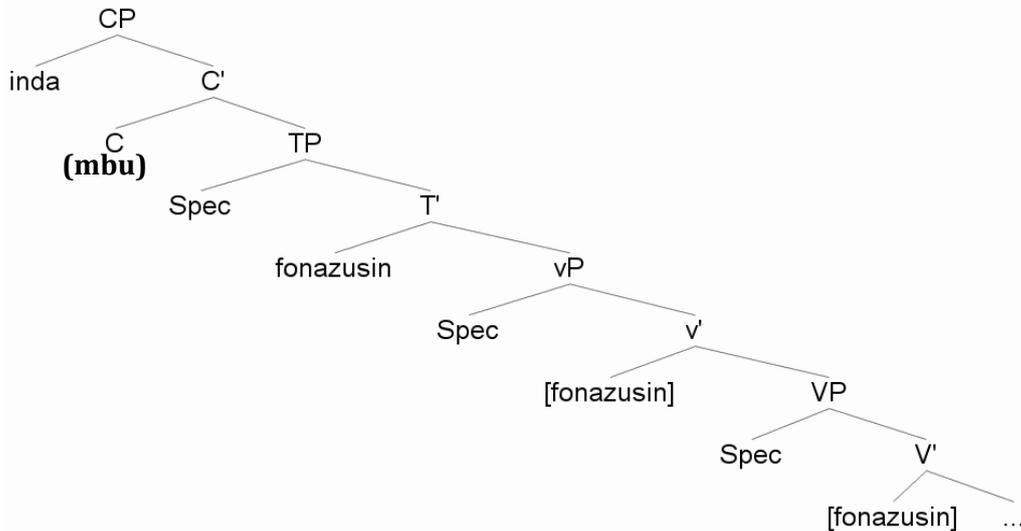
Based on the morphological properties of *inda mbu*, in *inda mbu* ‘what’, *mbu* is merged in C⁰ and *inda*, as the *wh*-phrase, is merged at Spec, CP. In this case, as has been observed in many languages, a *wh*-element can co-occur with an element in C⁰ contrary to the “doubly-filled COMP” (Chomsky and Lasnik 1977). Merging *mbu* at C⁰, as will be explained in more details below, follows from the need of a unified structure for both *mbu*-allomorphs and the variable *inda mbu*. As will be argued further on, *mbu* is on C⁰ because of the morphological properties of the allomorphs and the property of *inda* combining with an N in a complex *wh*-phrase, as mentioned in (2a), repeated here as (25). If *inda* can serve as one lexical item meaning ‘what’ when combined with an N, then it follows that the structure for *inda mbu* serving as an argument would look like (24).

- (25)a. *Inda fain* {*embu*, **mbu*} *emairepses?*
 What food.ACC *embu* cooked.2SG
 ‘What food did you cook?’

If indeed *mbu* is a complementizer, then following literature in D-linked *wh*-phrases, it should be ungrammatical when a *wh*-phrase ‘what’ is fronted with an overt Complementizer. Grewendorf (2008) in his attempt to explain ‘doubly filled COMP’ in Bavarian German lists *wh*-phrases in a linear order according to their operator-status, ranging from ‘why’ as the lowest one to ‘what’, as the highest one. He makes the generalization that the higher the degree of the operator of a *wh*-element, the lower the degree of grammaticality will be when it co-occurs with complementizer ‘that’. If we take this generalization to hold for complementizers other than ‘that’, it follows that the structure given in (24) should crash. But the lexical *wh*-phrase is argued to be here *inda*, which as mentioned in previous section can stand alone meaning ‘why’, and ‘why’ as argued by Grewendorf has a low degree of operator-status in D-linking. Further, as mentioned above, there is no clear indication related to the nature of *inda* for now rather just a simple presentation here as a *wh*-phrase in the specifier of CP.

Based on the morphosyntactic differences described in section 2 and following general distinction of the merging point of *wh*-adjuncts in the literature, *inda* is immediately merged in Spec, CP when it appears as stand-alone and means ‘why’.

(26) Inda (mbu) fonazusin?
 Why mbu shout.3PL
 'Why are they shouting?'

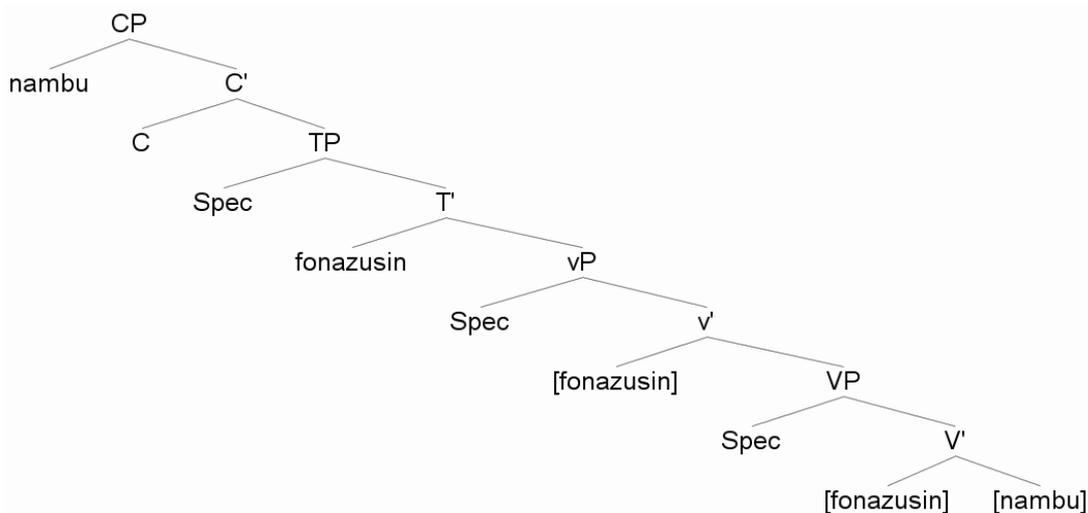


5.2. The three scenarios

5.2.1. The lexical scenario

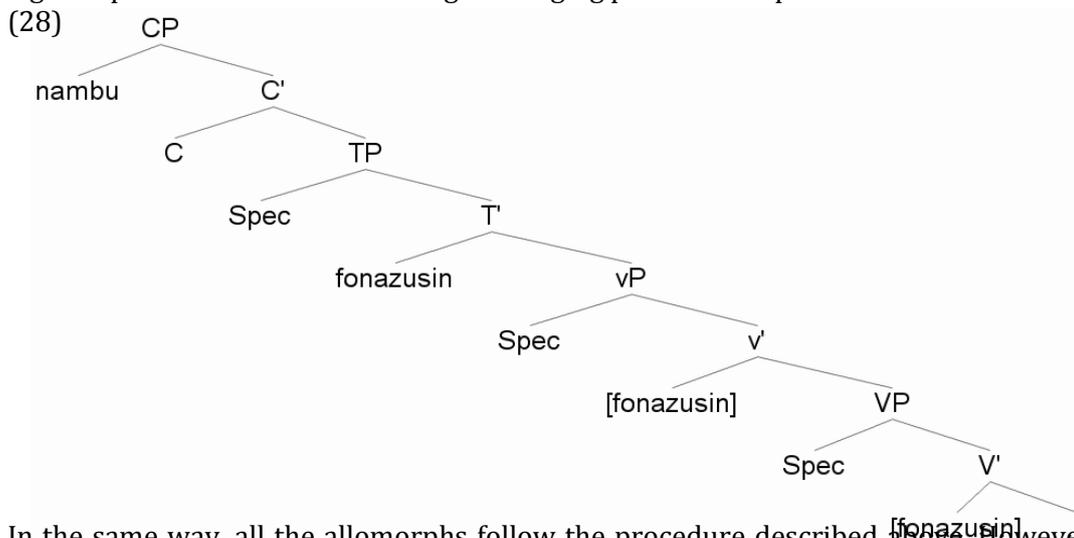
The *mbu*-allomorphs, as new items in the language, would be very logically entertained to be different lexical items that now exist in the lexicon. This would imply that the language change discussed above, as possible reason for their appearance is lexical and not grammatical. The status of these new items is that they are used as *wh*-questions and therefore should exist in the Spec, CP, as illustrated below for the example (20), repeated here as (27):

(27) a. Innambu/Nambu/Tambu/Ambu fonazusin?
 What/Why shout.3PL



Keeping in mind that Cypriot Greek is a null-subject language, the subject of the sentence can be omitted and therefore the order of *wh*-questions can be *nambu fonazusin (tsini)* 'What are they shouting', with the verb in T⁰. Agouraki (1997, 2001) argues that the verb in Cypriot Greek is at C⁰, except when C⁰ or Spec, CP is already filled. Following Chomsky's (1995) Copy Theory of Movement, *nambu*, as the internal argument, merges with the verb *fonazusin*. The original *nambu* is deleted and the copy of *nambu* is then merged to Spec, CP.

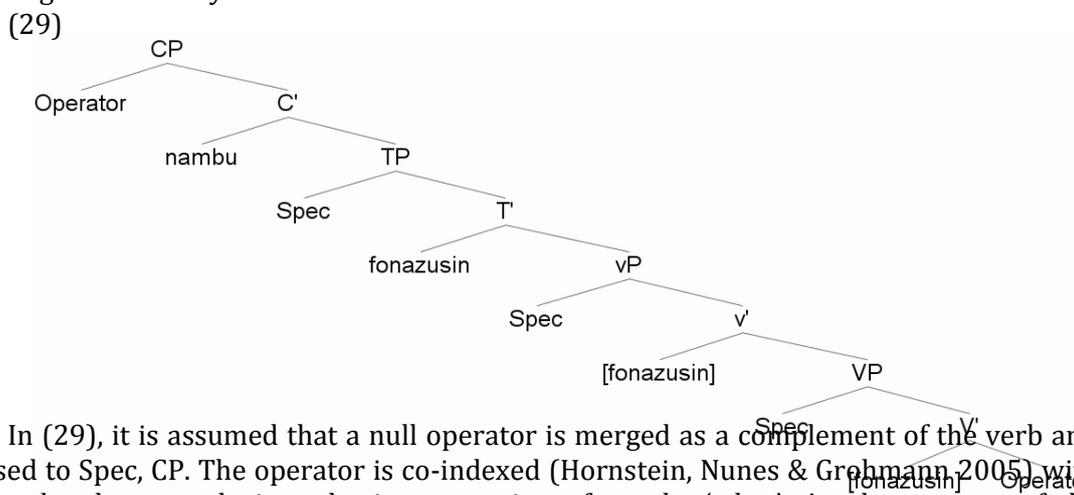
As mentioned in section 2, *nambu* can also serve as *wh*-adjunct. Assuming that adjuncts are merged directly in Spec, CP, this scenario leads to the standard assumption of having the specifier of CP as the landing or merging point for *wh*-phrases.



In the same way, all the allomorphs follow the procedure described above. However, there are some problems with this idea that need to be pointed out. *Innambu*, *nambu*, *tambu* and *ambu* can mean both ‘why’ and ‘what’. By saying that these allomorphs just like *inda mbu* (*wh*-argument) and *inda* (*mbu*) (*wh*-adjunct) are lexical items that exist independently in the lexicon of the speaker, then we immediately assume that there are two of each kind: an *innambu* meaning ‘what,’ an *innambu* meaning ‘why’, a *nambu* meaning ‘what’ and a *nambu* meaning ‘why’ etc. Indeed, the lexicon can be argued to be non-minimalistic for its containments but it is rather unnecessary to assume that we have the *mbu*-allomorphs, the variable *inda mbu* and possibly even the Greek *wh*-phrases *jiati* ‘why’ and *ti* ‘what’ because of the use of Standard Modern Greek on the island. Although nothing can be excluded, it is rather not economic and opposing to the Minimalist thinking to assume such an analysis for elements that show so similar properties. Considering their unifying properties of morphological difference with *inda mbu*, which sets them as one element with *mbu*, it is indeed easier to assume that they are lexical elements which are reinforced by the ongoing language change. But a minimalistic approach to the grammar rules out this analysis.

5.2.2. The operator-scope approach scenario

A second possible analysis for the *mbu*-allomorphs would be another possible landing site that they can be found:



In (29), it is assumed that a null operator is merged as a complement of the verb and raised to Spec, CP. The operator is co-indexed (Hornstein, Nunes & Grohmann 2005) with the *mbu*-phrase and gives the interpretation of *nambu* ‘what’. A relevant part of the literature deals with C⁰ in Cypriot Greek showing that it has a clause-typing feature that

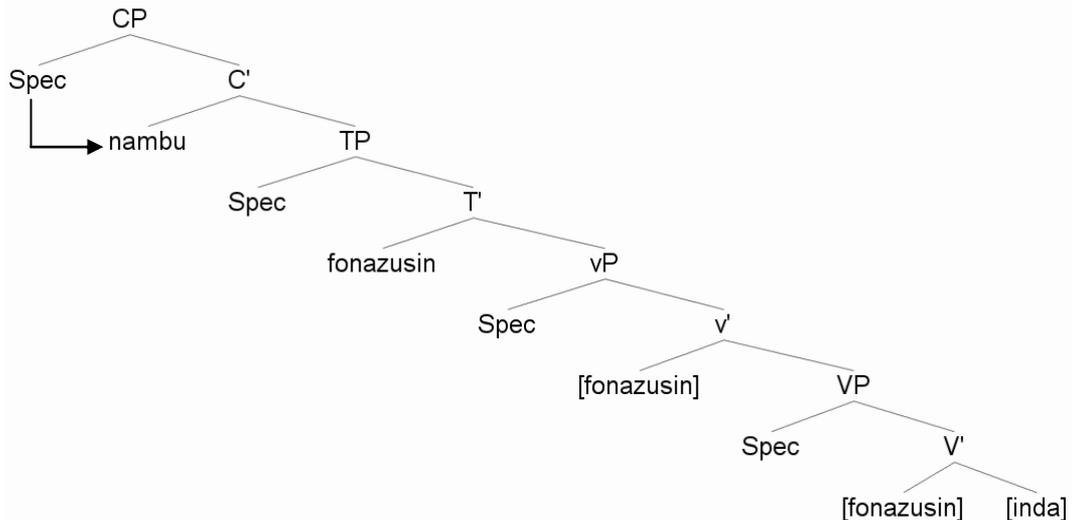
must be checked in the syntax (Agouraki 1997, 2001). Agouraki argues that this feature can be either negation raising to C^0 or a kind of Complementizer or a V-to-C rising. A possible reason for moving to C^0 in these cases, as she argues, is this feature since there has been already an operator, which is a preverbal stressed element and has filled the Specifier of CP. In her paper, she proposes that Cypriot Greek has a filled C requirement, referring specifically to the sentential force that needs to be checked overtly in C. In relevance to question-formation, there can be a specification [Question] in C, which is interpreted by the *wh*-questions in Spec, CP.

As mentioned above, Papadopoulou (in progress) claims that the Cypriot expression *mbu* in *wh*-questions is actually a complementizer found in C^0 .

Given that and following the same reasoning with Agouraki's claims, it can be assumed that there is some kind of operator in Spec, CP and that the *mbu*-allomorphs are elements in C^0 . Arguing that the allomorphs are indeed lexical items, there can be the case that *mbu* is actually an element targeting C^0 as Papadopoulou argues for *mbu*. Now, the problem appears to be that the *mbu*-adjuncts are supposed to be merged directly to C^0 , since Spec, CP is already filled by some kind of operator. This not only opposes to the distinction between true adjuncts and *wh*-arguments for merging in Spec, CP but also creates a problem since *wh*-adjuncts can merge into projections and not heads and implies that the problem is similar to the first scenario, leaving no space for explaining the difference between the *mbu*-arguments and *mbu*-adjuncts.

5.2.3. The lowering scenario

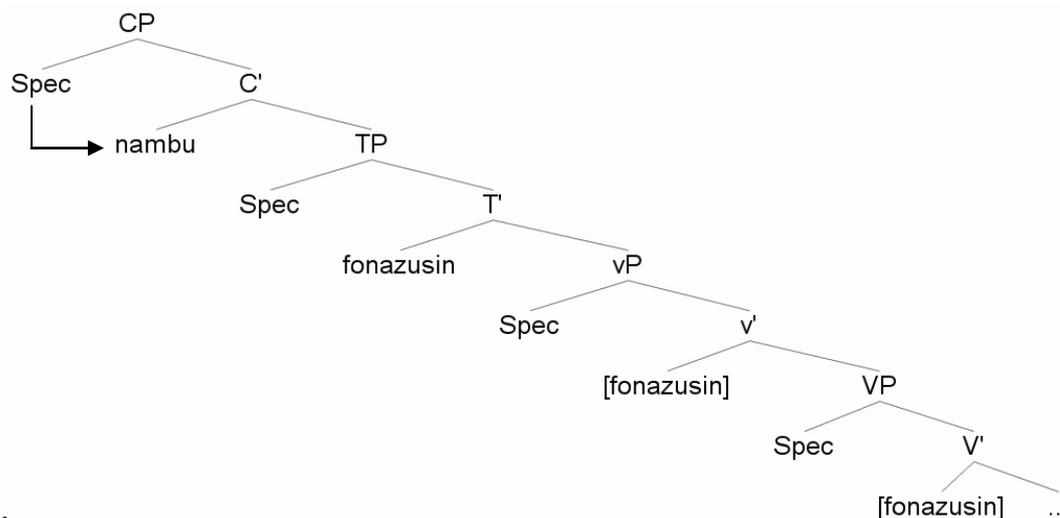
A third proposed scenario would be related to the previous one, namely that *mbu* needs to fill C^0 , but that does not mean necessarily that *innambu*, *nambu*, *tambu* and *ambu* are lexical elements which are copied there:



Mbu can exist on its own and *inna*, *na*, *ta* and *a* which are called to be possible allomorphs of the variable *inda* exist as one element which is the initial Cypriot *wh*-phrase before its changing; namely, *inda*. *Inda* is merged as the complement of the verb and then copied and remerged to Spec, CP. When our derivation reaches the projection of CP, *mbu* is merged in C^0 . Because *mbu* seems to be a strong element in syntax of Cypriot Greek based on all the properties examined so far (see section 2), it attracts the *wh*-phrase in Spec, CP and lowers it down to C^0 , so that it can be checked as one element that looks like *nambu* etc. Due to this attraction there are phonological processes coming in which turn the initial *inda* to *inna*- (when found with a topicalized element), *na*-, *ta*- and *a*-. These phonological or syntactic processes can be either called adjacency or fossilization (Papadopoulou in progress), hopefully to be explained clearer in the future. This would lead to the conclusion that the language change observed is not really an add of new elements in the lexicon but a grammatical change occurring in a syntactic and phonological level, namely

the function of *mbu* attracting *inda* and appearing as unifying elements i.e. *nambu* and not *na mbu*. It follows that a change in a morphosyntactic level can be argued to imply two things: To have as later implications, adaptation of Standard Modern Greek grammar, or the exact opposite which is that CG is in a completely different path than Standard Modern Greek.

(31)



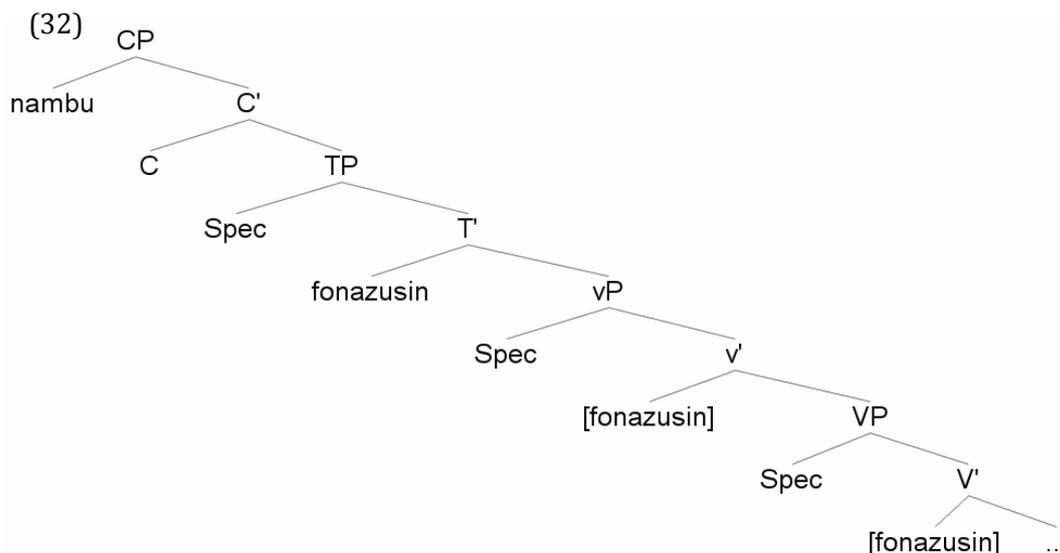
Adopting this scenario to *mbu*-adjuncts, the procedure is slightly changing. *mbu* is again an element which is merged directly to C⁰, but *inda*, merges directly to Spec, CP following again fundamental distinction on *wh*-arguments and true adjuncts. Then phonological processes and the strength of *mbu*, change *inda* to *inna-*, *na-*, *ta-* and *a-* and send it to LF as a unifying element.

This section discussed three possible analyses for the structure of *mbu*-allomorphs in the syntax. The first and second scenarios face the same problem: anti-economy! Assuming that new elements in language are lexical items only creates a lexicon with the *mbu*-allomorphs taking much more space than the theory accounts for. The lexicon can be by its nature not economic but, the ambiguities and the difficulty in processing the *mbu*-allomorphs as ‘why’ or ‘what’ imply that there are syntactic differences between the two. The second solution provided creates another problem, if one is to follow distinction between *wh*-adjuncts and *wh*-arguments. Having the *mbu*-allomorphs in C⁰, there is no merging point for adjuncts, but it assumes that either *mbu*-adjuncts exist as the *mbu*-arguments in the lexicon, which is excluded from the very start, or that they actually merge on C⁰. The third scenario places *mbu* in C⁰, and gives an analysis which is much closer to the real data than the other two. The similarity between *inda mbu* and its allomorphs *innambu*, *nambu*, *tambu* and *ambu* also leave strong implications for phonological processes.

5.3. The ambiguity in *mbu*-adjuncts and *mbu*-arguments

The three analyses given above examine various possibilities for the structure of *mbu*-allomorphs in the CG grammar but fail to account for the ambiguity between the *mbu*-adjuncts and the *mbu*-arguments. It is possible to think of the *mbu*-arguments following the third scenario and “blame” phonology for their unifying properties but it is not clear to say that *wh*-adjuncts follow the same procedure, too.

Following the distinction between *wh*-adjuncts and *wh*-arguments, then there can be only one solution left to be explored: *Mbu* is built up in the structure and is combined with *inda* to form the allomorphs, as analyzed in the previous section. One possibility is that the allomorphs which serve as adjuncts follow *inda* and are lexical items:



One of the main arguments following this analysis is that *inda* ‘why’, which is the reduced form if *inda mbu* ‘why/what’, exists as a lexical item in the grammar. In the process of language change, there can only be assumed its possible death but at the same time its replacement by the new elements. If *inda mbu* ‘what’ has a structure like (24) and the allomorphs follow the same pattern along those lines, then the allomorphs meaning ‘why’ can follow *inda* ‘why’ in (32).

However, as presented in Section 2, possible counter-arguments to this is that *inda* ‘what’ does exist as stand alone in some minorities in Cyprus (24a) or as a frozen expressions (24b) in the Cypriot population generally, and under this reasoning all the allomorphs should be lexical items. This possibility is already ruled out.

- (33)a. *Inda mairefkis?*
 What cooking.2SG
 ‘What are you doing?’
 b. *Inda kori?*
 Inda girl.NOM
 ‘What’s up girl?’

The reason for *inda* lacking a universal property of *wh*-phrases -like a stand-alone property- cannot be much explored by the analysis provided here. A possible reason is that *inda* is a fossilized element like *embu* (Papadopoulou in progress). If this is the case, the certain assumption is that this fossilization process, the change of *ine ti* ‘is what’ or *ti ine* ‘what is’ to a *wh*-phrase has absorbed any properties like stand-alone because of its once complex structure.

Other than that, it makes more sense for sentences like (20), repeated as (34), to have a different structure for *mbu*-arguments and *mbu*-adjuncts so as to get the difference in meaning. The difference in structure is simply assumed here to be of the different structural merging point of *wh*-objects and true adjuncts.

- (34) a. *Innambu/Nambu/Tambu/Ambu fonazusin?*
 What/Why shout.3PL
 #1 ‘What are they shouting?’
 #2 ‘Why are they shouting?’

5.4. A first restriction: *Innambu*

As discussed in section 4, the questionnaire was testing four syntactic environments, from which *innambu* seems to have a strong preference for use with topicalized elements. The semantic reason for the structural restrictions is not clear yet, but as can be inferred

from the speakers' comments it gives a stronger meaning, and thus gives emphasis to the topicalized element. Considering emphasis as the interpretation of *innambu*, it gives data to support Agouraki's claim (2008) on checking an [Emphasis] specification of sentential force on C of Cypriot Greek. The data that she gives have similar properties to the data of this questionnaire, and especially with the anaphoric form (for her the locative form *tsame* 'there' is here *dame* 'here') which has its interpretation to give some kind of emphatic meaning.

However, the difference between (35a) and (35b), is that in (35a) TSAME gives a contrastive meaning and it is thus argued to be a stressed element. In (35b), *dame* is referring to *tutos o mitsis*, which is treated as a topic of the sentence. So, *dame* forms one constituent with *tutos o mitsis* which is placed on Top⁰.

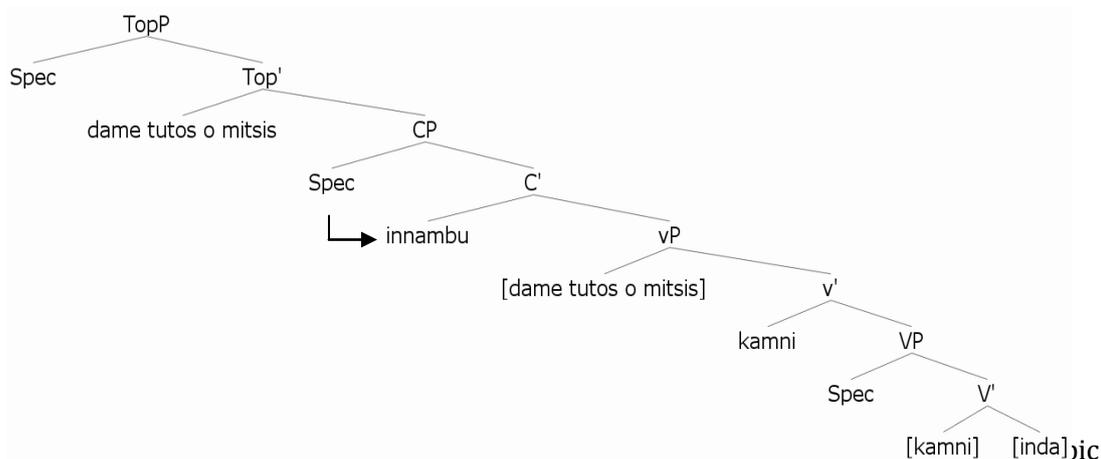
The syntactic distribution and the semantic contribution are not clear yet for *innambu* or any of the other allomorphs, since they are elements currently entering the language.

- (35) a. TSAME epia tse yo
 THERE went-1SG and I-NOM
 "I went just there/to the same place myself." (Agouraki 2008)
- b. Dame tutos o mitsis innambu kamni
 Here this the boy what do.3SG
 "What is this boy doing here?"

(Data in the questionnaire)

However, the difference in topic elements from stressed elements suggests a structure where there is a topic projection for this phrase. As mentioned before Spec, CP is already filled with the *wh*-phrase moved (*wh*-arguments) or merged (*wh*-adjuncts), so under this approach, even though *mbu*-phrases are in C⁰ (*wh*-arguments), Spec, CP cannot take any preverbal elements

(36)



element in the sentence and marks some kind of emphasis while pronouncing. If this is the case and based on the informative nature of the topics given, then there might be some relevance to the information focus. As has been inferred by speakers, emphasizing the topicalized element in *mbu*-questions gives difference in meaning as illustrated below for the sentence (37):

- (37)a. I thkyo tus innambu fonazun?
 The two.NOM them.POSS what shout.3SG
 "Why are the two of them shouting?"
- b. I **THKYO TUS** innambu fonazun
 The two.NOM them.POSS why shout.3SG

“Why are they shouting?”

It follows that the preference in syntactic environment with topicalized elements in the case of *innambu*, appears to have some relevance to the interpretation of allomorph ‘what’ and allomorph ‘why’. It is expected in future studies to see similar patterns and even more clearer restrictions for all the allomorphs discussed so far.

6. Conclusion

This paper discussed four new elements in the grammar of Cypriot Greek, which appear to be allomorphs of the standard form of the dialectal phrase *inda mbu*. A first comparison of *mbu* to *embu*, a Complementizer as argued by Papadopoulou (in progress) and a much more complex element according to Grohmann, Panagiotidis and Tsipaloku (2006) showed that the two show significant difference in their syntactic distribution. *Mbu* can only accompany *inda* serving as a *wh*-object or true adjunct, whereas *embu* cannot occur with *wh*-phrases functioning as the aforementioned *inda mbu*.

The four allomorphs of *inda mbu* appear to follow the same path, but differ in a morphosyntactic level. Their morphological properties are very much restricted compared to *inda mbu*, since they appear to behave as one element. Through a corpus selected by a questionnaire testing the four allomorphs in four possible syntactic environments produced by 100 speakers, it has been shown that there are some tendencies for a syntactic restriction in one of the allomorphs, the *innambu*, which appears to be preferred with a topicalized element. The morphosyntactic differences that appear for the allomorphs are argued to be the immediate result of ongoing language change observed in the corpus collected. The graphs given present an idiosyncratic pattern of language change, increasing the use of *nambu* in the youngest age group tested. The case of *innambu* shows a rare pattern of increasing tendencies of use in the age groups of 45-60 and 18-30. The corpus collected was restricted in the region of Limassol leading to the conclusion that *tambu* and *ambu* are allomorphs used in other regions of Cyprus, even though there has been a small number of data collected that show similar tendencies to *nambu* and *innambu*.

The existence of these four allomorphs in the grammar creates a question of their syntactic properties as *wh*-phrases. Having shown some tendencies characterized by different frequency of use, it is still not clear whether these differences in frequency will be eliminated once language change has been completed. If not, then there should be a reason following current syntactic work accounting for the co-existence of the allomorphs and their use by speakers independently of any external factors. Based on the data collected, a syntactic approach which accepts the allomorphs as lexical forms in the lexicon is ruled out, since it does not account for any semantic difference but created a number of *mbus* in the lexicon. Following relevant work on syntactic approaches to Cypriot Greek (Agouraki 2008), the second scenario excludes the possibility of accepting the allomorphs as lexical elements which target C^0 . The use of a null operator in Spec, CP co-indexed with the *wh*-phrase in C^0 creates problems in arguing that *mbu*-adjuncts merge immediately to Spec, CP whereas *mbu*-objects are copied after merged with the verb. A last suggestion puts *mbu* in C^0 and presupposes that the initial form of the allomorphs is *inda*, which after merged with the verb and copied to Spec, CP is attracted by *mbu* and lowers down to C^0 changing in *na-*, *inna-*, *ta-* and *a-* due to phonological processes.

These newly-appeared allomorphs in CG contribute to the discussion of *wh*-questions, the relevance of the overt complementizers and the possible function of them as one element (Papadopoulou in progress) or deconstruction of them as clefts, as argued for *embu* (Grohmann, Panagiotidis & Tsipaloku 2006). The phenomenon of their unifying properties is yet syntactically and phonologically undetermined, but this paper offers the most significant properties characterizing them. Future work (Pavlou in progress) concentrates on the nature of *inda*, presented here as a *wh*-phrase, and its possible decomposing as a cleft in its combination with *mbu*. In relevance to this and in addition to

the already existed corpus of the acquisition of *wh*-phrases and relevant structures in CG (Papadopoulou in progress), it is aimed that the acquisition of the structures listed here will be tested from their acquisition perspective.

References

- Abney S. (1987). *The English Noun Phrase in Its Sentential Aspect*. PhD dissertation, Massachusetts Institute of Technology, Cambridge.
- Agouraki Y. (1997). On the Enclisis/ Proclisis Alternation. , in Drachman, G., Malikouti-Drachman, A., Fykias, J., Klidi, C. (eds.) *Proceedings of the Second International Conference on Greek Linguistics*, Salzburg: University of Salzburg Vol. II: 393-404.
- Agouraki Y. (2001). The position of clitics in Cypriot Greek, in Ralli, A., Joseph, B., Janse, M. (eds.), *Proceedings of the First International Conference of Modern Greek Dialects and Linguistic Theory*, Patras: University of Patras, 1-18.
- Agouraki Y. (2010). It-clefts and Stressed Operators in the Preverbal Field of Cypriot Greek. *Lingua* Vol. 120/3: 459-780.
- Bailey G. (2002). Real and Apparent Time. *The Handbook of Language Variation and Change*. UK: Blackwell Publishing Ltd.
- Chambers J.K. (1994). An Introduction to Dialect Topography. *English World-Wide* 15: 35-53.
- Chambers, J.K. & Trudgill P. (1998). *Dialectology*. 2nd ed. Cambridge: Cambridge University Press.
- Chambers J.K. (2002). Patterns of Variation Including Change. *The Handbook of Language Variation and Change*. UK: Blackwell Publishing Ltd.
- Chomsky N. *The Minimalist Program*. Cambridge, MA:MIT Press, 1995.
- Chomsky N. & Lasnik H. (1977). Filters and Control. *Linguistic Inquiry* 8: 425-504.
- Fotiou K. (2009). Focusing strategies in Cypriot Greek. In K.K. Grohmann & P. Panagiotidis (eds.), *Selected Papers from the 2006 Cyprus Syntaxfest*, Newcastle upon Tyne: Cambridge Scholars Publishing, 63-91.
- Grewendorf G. (2008). *Wh-movement as topic movement*. Mskr. Universität Frankfurt a.M.
- Grohmann K. (to appear). Issues in and for the acquisition of Cypriot Greek. In E. Rinke & T. Kupisch (eds.) *The Development of Grammar: Language Acquisition and Diachronic Change*. Amsterdam: John Benjamins.
- Grohmann K., Panagiotidis P. & Tsiplakou T. (2006). Properties of *Wh*-Question Formation in Cypriot Greek. In M. Janse, B. D. Joseph and A. Ralli (eds.) *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory (Mytilene, Greece: 30 September-3 October 2004)*. Patras: University of Patras, 83-98.
- Grohmann K.,Theodorou E., Pavlou N. & Leivada E. (2010). Object Clitics in Typically- Developing and Language- Impaired Cypriot Greek Children. *Let the Children Speak. Learning of Critical Skills Across 25 languages*, London (22-24 January, 2010).
- Grohmann, K. & Papadopoulou E. (to appear). Question(able) issues in Cypriot Greek, *Linguistic Analysis* 35.
- Henry, A (2002) Variation and Syntactic Theory. *The Handbook of Language Variation and Change*. UK: Blackwell Publishing Ltd.
- Hockett C. F. (1958). *A Course in Modern Linguistics*, New York: Macmillan.
- Hornstein N., Nunes J. & Grohmann K. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.
- Ioannidou A. & Pavlou N. (2009). *The Awareness of Being Aware: The Case of Cyprus*, unpublished ms., University of Cyprus, Nicosia.
- Labov W. (1963) The social motivation of a sound change. *Word* 19: 273-309.
- Labov W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov W. (1982). *The social stratification of English in New York City*. Washington: Center for Applied Linguistics.
- Montgomery, M. (1997). A Tale of Two Georges. In J. Kallen (ed.) *Focus on Ireland*. Amsterdam: Benjamins.
- Panagiotidis, P. (2009). *On the dissociation between syntax and morphology: evidence from 'mixing' in Cypriot Greek*. Paper presented at the 9th International Conference of Greek Linguistics, University of Chicago (October 29-31, 2009).
- Papadopoulou, E. (in progress). *Acquisition of wh-questions: Evidence from Cypriot Greek*. Essex University: PhD dissertation.

- Pavlou N. (in progress). *Inda mbu or ine ti pu?* The cleft-strategy in Cypriot Greek *wh*-phrases. Unpublished ms., University of Cyprus, Nicosia.
- Simeonidis, C. P. (2006). *Ιστορία της Κυπριακής Διαλέκτου*. Nicosia: Holy Monastery of Kykkos Centre of Research.
- Spyropoulos, V. (1999). The structure of small clauses in Modern Greek. In Brian D, Joseph, Geoffrey C. Horrocks & Irene Philippaki-Warbuton (eds.) *Themes in Greek Linguistics II*: 169-196.
- Tsai, W.-T. D. (2008). Left Periphery and *How-Why* Alternations. *Journal of East Asian Linguistic* 17: 83-115.
- Wang, W. & Cheng C.-C. (1970). Implementation of phonological change: The Shûang-fêng case., *Chicago Linguistic Society* 6: 552-527.

Ορισμένες σκέψεις σχετικά με την προσωδία του Γκρίκου

ANTONIO ROMANO

ΦΡΑΝΤΣΕΣΚΑ ΠΑΠΑΣΠΥΡΟΥ

ΠΑΥΛΟ ΜΑΪΡΑΝΟ

Τμήμα Γλωσσικών Επιστημών – Σχολή Ξενόγλωσσων Φιλολογιών

Πανεπιστήμιο του Τορίνο, Ιταλία

antonio.romano@unito.it, papaspiru@libero.it, paolo.mairano@gmail.com

Abstract

The aim of this paper is to illustrate some basic intonational properties of dialogues in the Greek dialects spoken in the South-East region of Italy called *Grecia Salentina*.

As has already been noted in previous studies, speakers of this region form a multi-lingual community and are able to switch among at least three interfering codes: a (more or less regional) variety of Italian, a Romance dialect (akin to other Sallentinian dialects found in the surrounding area), a Greek dialect (which is gradually being abandoned). In recent years, Modern Greek – mainly in its written form – has gained a limited diffusion as a fourth interfering code for a few speakers (socio-cultural aspects related to this phenomenon have been discussed in [21]).

One of the characteristics of this Greek dialect (now called *griko*) that have been more thoroughly analysed is lexis, which has been studied with the help of dictionaries and text collections published in the last century and onwards (e.g. [12]; see comprehensive surveys in [4, 5, 6, 13]). Moreover, a few studies on specific syntactic properties began to appear – sometimes carried out with experimental methods (see [11]; also cp. [22] and [8, 9], for Greek spoken in Calabria) – as well as a number of grammars.

Apart from a few isolated remarks about macroscopic interdialect phonetic differences, phonological features have been the least studied properties as they have usually been judged of secondary importance above all where they are not connected with the development of spelling conventions (nevertheless, we find important exceptions such as [14, 15] and [10]). A few studies deal with syllable structure, consonant gemination and acoustic properties of vowels ([18, 19]), but if one excludes stress (mainly discussed in [16]), prosodic features are usually neglected: important prosodic phenomena such as rhythm and intonation still remain undescribed (see [17]).

The topic is very relevant if one considers the general convergence/divergence dynamics between prosodic systems of different languages in contact; particularly, data may shed light on when and how the simultaneous use of these codes was established within the centuries.

Yet, this kind of data needs to be collected and scrutinised in very specific and peculiar modes. Objective interpretations are only possible within rigorous comparison schemes needing a high degree of sophistication which cannot be reached outside an advanced analytic framework, which is usually unavailable to local researchers. Moreover, it is difficult to collect spontaneous utterances if one considers that suitable fluency conditions are restricted to elderly speakers, whereas hybrid productions (caused by the interference of the other codes) are nowadays common in younger people.

By following an analytic approach already tested in other scientific domains, it has been possible to carry out an intonational analysis on the controlled recordings provided by [23]. Among the more stable intonation contours which were isolated we found typical profiles (well distinct from the ones discussed by [1, 2, 3, 7]) which are however shared by at least the two dialectal varieties (*Griko* and Romance Sallentinian).

1. Εισαγωγή

Αυτό το άρθρο στοχεύει να διασαφηνίσει κάποια σημαντικά προσωδιακά χαρακτηριστικά ορισμένων ελληνικών διαλέκτων, που μιλιούνται, μέχρι σήμερα, στο νοτιοανατολικό τμήμα μιας Ιταλικής περιοχής, γνωστή με το όνομα *Grecia Salentina*.

Όπως έχει αποδειχθεί, από προηγούμενες μελέτες, η γλωσσική κατάσταση αυτής της περιοχής χαρακτηρίζεται από την παρουσία πολύγλωσσων κοινοτήτων, όπου οι ομιλητές εναλλάσσουν τουλάχιστον τρεις γλωσσικούς κώδικες: έναν ιταλικό (σχεδόν τοπικό), μία ρωμανική διάλεκτο (η οποία έχει σημαντικές ομοιότητες με άλλες διαλέκτους της περιοχής του Σαλέντου, όπως και αρκετά πρωτότυπα στοιχεία), μία ελληνική διάλεκτο, η οποία ομιλείται πάντα λιγότερο (της οποίας η προέλευση συζητήθηκε πολύ στο παρελθόν

εις βάρος των συγχρονικών σπουδών που θα έδιναν μεγαλύτερη αξία στα ζωτικά χαρακτηριστικά και ίσως θα είχαν βοηθήσει την διατήρηση).

Σχετικά με την σποραδική παρουσία της προφορικής νεοελληνικής γλώσσας (που από καιρό κυκλοφορεί και σε γραπτή μορφή), στην παραγωγή λόγιων ομιλητών, μπορεί να συζητηθεί ξεκινώντας από τα δεδομένα που παρουσίασαν ο [21].

Μεταξύ των χαρακτηριστικών αυτής της διαλέκτου, που κάποτε ονομαζόταν *greco otrantino* (ελληνικά του Ότραντου) ενώ σήμερα είναι γνωστή ως *griko*, έχουν μελετηθεί κυρίως τα λεξιλογικά χαρακτηριστικά (με εμβάθυνση από πλευράς πολλών λεξικών και δημοσιεύσεων του περασμένου αιώνα, βλέπε συνοπτικές έρευνες [4, 5, 6, 13]), έστω κι αν δεν απουσιάζουν μελέτες σχετικά με τις συντακτικές ιδιότητες της γλώσσας, για τις οποίες, κάποιες φορές, χρησιμοποιήθηκαν και πειραματικές μέθοδοι (βλ. [11], και [8, 9], για τα ελληνικά της Καλαβρίας, βλ. [22]).

Τα στοιχεία που έχουν μελετηθεί λιγότερο, εκτός όταν η ανάγκη οδηγεί στην υιοθέτηση μιας κοινής γραφής, είναι αναμφίβολα αυτά που σχετίζονται με τις φωνητικές ιδιότητες (όμως και σε αυτήν την περίπτωση δεν λείπουν οι εξαιρέσεις, βλ. [14, 15] και [10], για τα ελληνικά στην Καλαβρία, και άλλες αναφορές στη βιβλιογραφία), κυρίως από διαλεκτολογική και κοινωνικογλωσσική άποψη η οποία στοχεύει στην παρατήρηση της κοινωνικογεωγραφικής μεταβολής.

Τέλος έχουν, ιδιαίτερα, παραμεληθεί οι έρευνες σχετικά με τις προσωδιακές απόψεις αυτών των γλωσσικών ποικιλιών. Αν εξαιρεθούν οι σποραδικές αναφορές σχετικά με τον τονισμό των λέξεων (κυρίως του [16]) και ελάχιστες μελέτες σχετικά με τη συλλαβική δομή και το διπλασιασμό των συμφώνων (π.χ. [18, 19]), καταλήγουμε στο συμπέρασμα ότι, τα σημαντικά προσωδιακά φαινόμενα, όπως, ο ρυθμός και ο επιτονισμός παραμένουν ακόμη ευρέως άγνωστα.

Το θέμα γίνεται πολύ πιο ενδιαφέρον, αν θεωρηθεί ότι η παρατήρηση των δυναμικών σύγκλισης και απόκλισης μεταξύ των προσωδιακών συστημάτων των διαφόρων κωδικών, που εμπλέκονται σε αυτές τις κοινότητες, θα μπορούσαν να εμφανίσουν καινούργιες δυνατότητες σχετικά με το χρόνο και τους τρόπους με τους οποίους καθιερώθηκαν οι συνθήκες επαφής μεταξύ αυτών των ομιλιών κατά τη διάρκεια της ιστορίας.

2. Υλικά και μέθοδος

Στοιχεία τέτοιου είδους έχουν, όμως, την ανάγκη να συλλεχθούν, να ταξινομηθούν και να συγκριθούν κατά έναν τρόπο που επιζητά υψηλό επίπεδο επιτήδευσης που, όμως, δύσκολα μπορεί να επιτευχθεί, εφόσον σήμερα πολλοί από τους ομιλούντες, που θα μπορούσαν να προσφέρουν μια αυθόρμητη παραγωγή λόγου είναι τώρα πια, μεγάλης ηλικίας ενώ μεταξύ των νεότερων είναι, πλέον, κοινό ότι μπορούμε να έχουμε μόνο λύσεις αμβισβητούμενες και / ή αλληλότυπες.

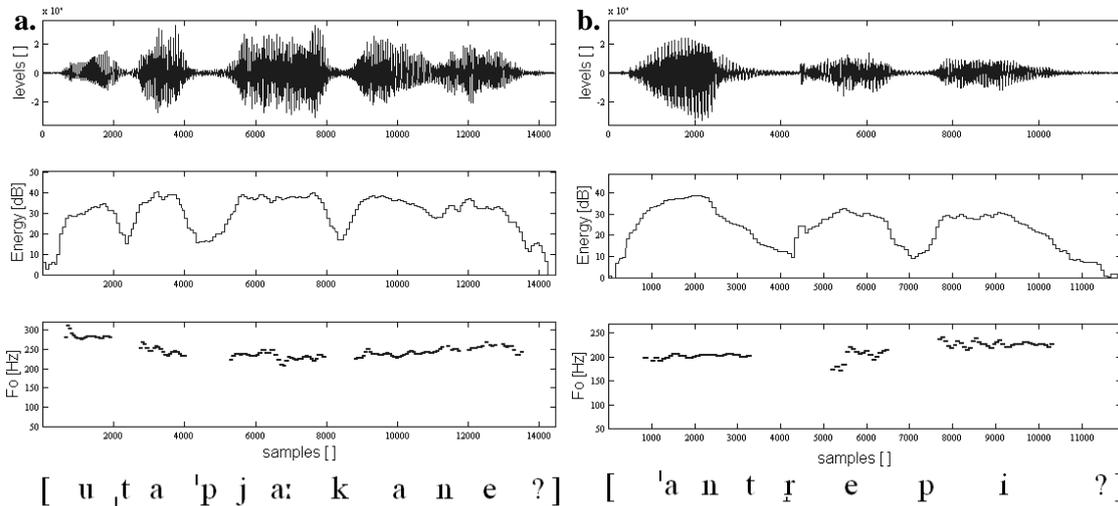
Παρόλο αυτά, ακολουθώντας μία αναλυτική πρόσβαση, που έχει ήδη πειραματιστεί σε άλλα θέματα, μπόρεσαν να γίνουν επιτονικές αναλύσεις πάνω σε πρωτότυπες καταγραφές και σε δημοσιευμένα στοιχεία του [23], οι οποίες έδωσαν σημαντικά αποτελέσματα: τα πιο χαρακτηριστικά σημεία της επιτονικής καμπύλης μοιάζουν να είναι κοινά τουλάχιστον για τις δύο διαλεκτικές ποικιλίες (*griko* και *salentino*).

Στο πλαίσιο αυτής της έρευνας, ακόμη σε στάδιο εξέλιξης, αναλύσαμε τα 50 μελωδικά σχήματα που συναντώνται στις πρώτες 22 παρεμβάσεις των 5 ομιλητών που συμμετείχαν στο διάλογο “To vikènde” (βλ. [23]: 24-29, που αφορά μια συνολική περιγραφή μερικών εξ αυτών, βλ. προσάρτημα).

3. Αποτελέσματα

Μερικά από τα προφίλ που βρίσκονται στο υλικό μας, έστω κι αν χαρακτηρίζονται από μια γνωστή εκφραστική εναλλαγή ύφους (το οποίο εξαρτάται από το είδος του κειμένου), είχαν ήδη επιλεγεί για την εξέταση ορισμένων κανόνων επιτονισμού και προσωδίας που συναντήθηκαν σε ένα *corpus fisso* (πειραματικού τύπου, βλ. [17]).

Ορισμένα από αυτά (όπως εκείνα στην Εικ. 1) δείχνουν την τάση, παρατηρήσιμη τουλάχιστον στην ποικιλία της πόλης Calimera, η οποία παρουσιάζει και το μεγαλύτερο ενδιαφέρον στην περιοχή της Lecce, ως προς την αναπαραγωγή ενός ολικού προσωδιακού ερωτηματικού σχήματος (*y/n question*) το οποίο εμφανίζει μια τυπική ισοπέδωση σε σχέση με τη συλλαβή πυρήνα, κατεβαίνοντας από τιμές ποικιλοτρόπως υψηλότερες, αλλά και ανεβαίνοντας προοδευτικά περίπου κατά μισό τόνο στις επόμενες άτονες συλλαβές, δημιουργώντας έτσι ένα χαρακτηριστικό τονικό προφίλ *untuned*.

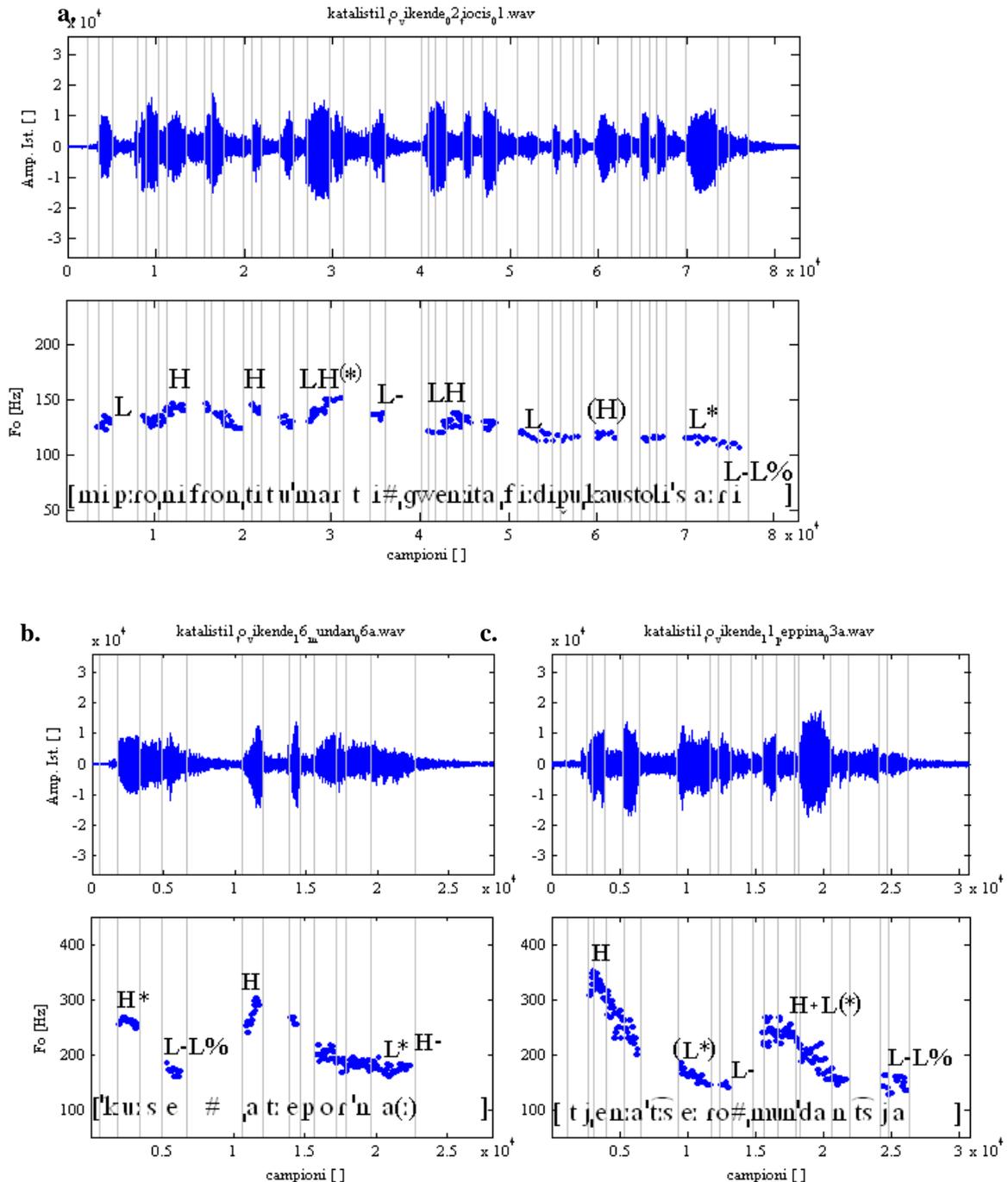


Εικόνα 1: *Waveform, loudness και pitch curves για δύο ερωτήσεις από δύο ομιλήτριες γυναίκες από την Calimera: a. MD59, u ta piàkane? "του το έφεραν;"*, b. PA69, àntrepi?! "άνδρες;!" (προσαρμογή από [17]).

Ένας κάποιος αριθμός προφίλ παρουσιάζει δυσκολία ταξινόμησης διότι υφίσταται εναλλαγές εμφατικού τύπου (ή πρόκειται για το αποτέλεσμα μιας υποκριτικής τέχνης). Ανάμεσα από αυτούς τους διαλόγους οι πιο διαδεδομένοι, εκτός από τους ερωτηματικούς, είναι στην ουσία οι δηλωτικοί (όπως εκείνος στο 02_01), οι αποσιωπητικοί (συνεχόμενοι, όπως στο 16_06a) ή οι κλητικοί (11_03a ή 19_06a, βλ. Εικόνα 2).

Εξαιρώντας τα σχήματα με αυτό το σκοπό, μια σημαντική ποσότητα των παρόντων προφίλ στο υλικό που παρουσιάζουμε εδώ, χαρακτηρίζεται από την ιδιαίτερη συχνότητα ενός εκθεσιακού ύφους κι όχι ουδέτερου, μονοτονικού στα προπυρηνικά τμήματα ορισμένων περιεχομένων και ιδιαίτερα φθίνοντας (ξεκινώντας από τη συλλαβή πυρήνα κι ως εκ τούτου χαμηλού στη επόμενη της επιτονικής συλλαβής) όταν πρόκειται για συμπερασματικό ύφος (στη διαβεβαίωση) ή αύξοντος στους υψηλά πυρηνικούς τόνους όταν πρόκειται για επίκληση.

Οι προσωδιακές αποστάσεις, που έχουν αξιολογηθεί μέχρι τώρα, μεταξύ αυτών των προφίλ και όσων αντιστοιχούν στο Salentino, υπογραμμίζουν τις συνθήκες πλήρους αναγνώρισης ανάμεσα στα δύο προσωδιακά συστήματα όσον αφορά τη δομή και την έκφραση (βλ. [17]). Αυτό γίνεται ακόμη πιο πραγματικό όταν τα χαρακτηριστικά προφίλ, που περιγράψαμε, χρησιμοποιηθούν για άλλες μορφές ελληνικών (βλ. [1, 2, 3, 7]).



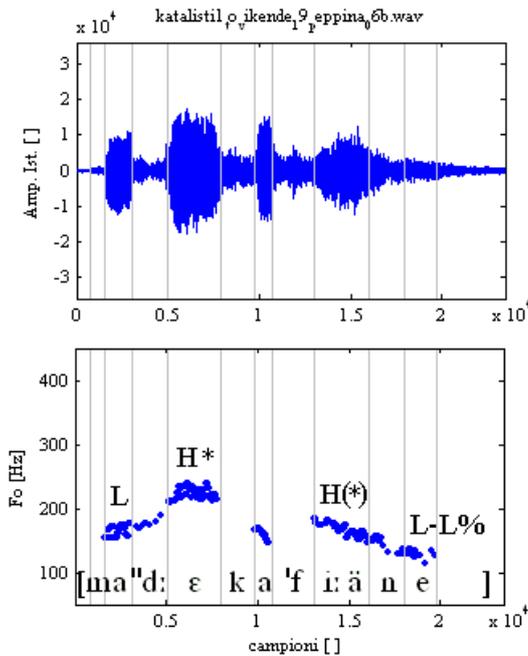
Εικόνα 2: Waveform, loudness και pitch curves για: a. 02_01, *Mi prroni fronti tu marti guenni t'afidi ru kau sto lisari*. “Με την πρώτη βροντή του Μάρτη βγαίνει το φίδι κάτω από την πέτρα”, b. 16_06a, (Kuse:) *atterornà*, “Άκουσε: σήμερα το πρωί”, c. 11_03a. (Ti ènna tzero,) *Mundantzia?* “Τι πρέπει να ξέρω, Μουντάντσια;”.

Σαν συμπέρασμα αυτών των σύντομων παρατηρήσεων, επιμένοντας στα ιδιαίτερα χαρακτηριστικά του *code mixing* που βρίσκονται σε αυτό το υλικό, εμφανή κυρίως σε επίπεδο προσωδίας αλλά παρόντα και σε λεξικολογικό επίπεδο (λεξικό και φρασεολογία), έτσι όπως δείχνουν οι ακόλουθες παρατηρήσεις.

Στους διαλόγους που έχουν αναλυθεί, παρατηρούνται πολυάριθμα στοιχεία λεξιλογικής παράλλαξης μεταξύ γρίκου, σαλεντίνου και ιταλικών (*annamurà* ‘innamorato’ ‘ερωτευμένος’, *frastornà* ‘frastornata’ ‘ζαλισμένη, χαμένη’, *mpalà*(n) < *mpalata* ‘impalata, immobile, ferma’ ‘ακίνητη’, *mpoggetzi/empoggei* < *mpuggiare* ‘appoggiare, riposarsi’ ‘ακουμπάω, ξεκουράζομαι’, *penserria* ‘pensieri’ ‘σκέψεις’, *paradiso*

'paradiso' 'παράδεισος, *skiattetzi* < *schianti* 'crepi' 'πέθανε', *sordu* < *sordi* 'soldi' 'λεφτά', *vekkia* 'vecchia' 'γριά, χειμώνας', *vòti(se)* < '(ti) volti' 'απευθύνεσαι') και για πιο λειτουργικά στοιχεία (*doru* 'doro' 'μετά', *kùkkia* 'vicina a' 'κοντά', *kundu* 'come' 'σαν', *largo* 'lontano' 'μακριά', *ma* 'μά'...). Πολυάριθμα είναι τα σημασιολογικά και μεταφραστικά δάνεια και στον τομέα των δεικτών ομιλίας (*discourse and focus markers* – *certo* 'certo' 'βέβαια', *pròbbio* 'proprio' 'ακριβώς', *puru* 'pure' 'επίσης', (*o*) *justo* 'giusto' 'σωστά', *però* 'però' 'όμως'... Για παρόμοια τα αποτελέσματα βλέπε [22]).

Σε επίπεδο καθαρά προσωδιακό επαληθεύτηκε ότι, στα κοινά σχήματα γενικής οργάνωσης που αναλύθηκαν σύμφωνα με τα προσωδήματα που προτάθηκαν από τον [20], συνδέονται τα εξειδικευμένα προφίλ που αντιστοιχούν στα χαρακτηριστικά του Σαλέντου της ίδιας της βοριοκεντρικής περιοχής (Iccese) (βλ. [17, 18]). Λύσεις εντοπισμού και/ή τεμαχισμού είναι, συχνά, ίδιες με αυτές που βρίσκονται στη διάλεκτο του Σαλέντου ακριβώς όπως συμβαίνει και με ορισμένους ιδιοματικούς τύπους: *Ka stèun oli mia!* 'Che stanno tutti una!' (reg. It. "stare tutt'una" 'είναι κανείς σύμφωνος') 'είναι όλοι σύμφωνοι!' και με κάποιες υφολογικές εκφράσεις που αναφέρονται στα κοινά κοινωνικο-πολιτιστικά στοιχεία (*addho ka ssordu!* 'altro che soldi!' 'κάθε άλλο παρά λεφτά!', *ka fidete na stasi kammeni arte e Peppina?!* – Sal. *ca se fide sse stescia ssettata mo' la Peppina?!* 'Che riesce a stare seduta adesso la Peppina?!' 'Θα τα καταφέρετε να καθήσει τώρα η Peppina!;', (*Tuo en jalissio,*) *Ma de' kka fiane!* – Sal. (*quistu è' bberu,*) *ma no' cca (se nde) fuscira!* '(questo è vero,) ma non che fuggirono!' ("fuga d'amore") '(αυτό είναι αλήθεια,) μα όχι πως φύγανε (πως κλεφτήκανε)' (βλ. Εικόνα 3).



Εικόνα 3: Waveform, loudness και pitch curves για 19_06b, *Ma de' kka fiane*. "Μα όχι πως φύγανε".

4. Συμπεράσματα

Σε αυτή την εργασία, ακόμη σε εξέλιξη, αρχίσαμε την πειραματική αξιολόγηση ποσότητας και ποιότητας των πιθανών αναλογιών μεταξύ προσωδιακών σχημάτων των διαφόρων γλωσσικών κωδίκων που υπάρχουν στις γλωσσικές παραγωγές των ομιλητών του Ελληνικού Σαλέντου. Εκτός από τα ειδικά στοιχεία σύγκλισης και απόκλισης μεταξύ των χαρακτηριστικών λύσεων ορισμένων εξετασθέντων κωδίκων, βρίσκεται σε εξέλιξη μια ποσοτική αξιολόγηση πιο συστηματική και με κριτήρια απόστασης.

Η προσωδιακή απόσταση είναι ιδιαίτερα χαμηλή σε δηλωτικές εκφωνήσεις, αλλά παρουσιάζει ενδιαφέρον όσον αφορά τους βαθμούς σύγκλισης και στις ερωτηματικές εκφωνήσεις, καθώς παρατηρείται μια κοινή γραμμή με τις ρωμανικές διαλέκτους των γειτονικών περιοχών.

Οπότε και από αυτή την άποψη ισχύουν όσα είπε η Μ. Κατσογιάννου ([8]: 516, σχετικά με άλλες απόψεις γύρω από τα ελληνικά της Καλαβρίας): η διάλεκτος *griko* βρίσκεται πολύ πιο κοντά στη διάλεκτο του Σαλέντου (*salentino*), παρά στα νεοελληνικά, ακριβώς όπως η διάλεκτος του Σαλέντου είναι πιο κοντά στη διάλεκτο *griko*, παρά στην ιταλική γλώσσα.

Βιβλιογραφικές αναφορές

- [1] Arvaniti A. & Ladd D. R. (1995). Tonal alignment and the representation of accentual targets, *Πρακτικά του "13th ICPhS"* (Stockholm, Σουηδία, 1995) 4: 220-223.
- [2] Botinis A. (1992). Accentual Distribution in Greek Discourse. *Travaux de l'Institut de Phonétique d'Aix* 14 : 13-52.
- [3] Botinis A. (1998). Intonation in Greek. In D. Hirst & A. Di Cristo (eds.) *Intonation Systems: a Survey of Twenty Languages*, 288-310. Cambridge: Cambridge Univ. Press.
- [4] Fanciullo F. (1994). Latino e greco nel Salento. In B. Vetere (ed.), *Storia di Lecce. Dai bizantini agli Aragonesi*: 421-486. Bari: Laterza.
- [5] Fanciullo F. (2001). On the origins of Modern Greek in Southern Italy. *Πρακτικά του 1ου Διεθνές Συνέδριου Νεοελληνικών Διαλέκτων και Γλωσσολογικής Θεωρίας* (Πάτρα, 2000), επιμ. Α. Ralli, B.D. Joseph & M. Janse, Πάτρα: Πανεπιστήμιο Πατρών, 67-78.
- [6] Fanciullo F. (2004). Incontri linguistici nel Medioevo mediterraneo. *Dialetti e non solo*, Alessandria: Dell'Orso, 57-74.
- [7] Fujisaki H., Ohno S. & Yagi T. (1997). Analysis and Modeling of Fundamental Frequency Contours of Greek Utterances. *Proceedings of EuroSpeech '97 (5th European Conference on Speech Comm. and Technology*, Ρόδος, 1997), 465-468.
- [8] Κατσογιάννου Μ. (1997). Interventi simbiotici tra greco e romanzo nell'area linguistica calabrese. *Proceedings of the "2o Incontro Internazionale di Linguistica greca"* (Τρέντο, Ιταλία, 1995), Univ. di Trento: Dip. di Scienze Filologiche e Storiche, 514-528.
- [9] Κατσογιάννου Μ. (1997). Ελληνικά της Κάτω Ιταλίας: μορφολογία των ονομάτων και εξέλιξη του κλιτικού συστήματος. *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 17ης Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας του Α.Π.Θ.* Θεσσαλονίκη: Αφοί Κυριακίδη, 328-341.
- [10] Κατσογιάννου Μ. (1999). Το ιδίωμα της Καλαβρίας. *Διαλεκτικοί θύλακοι της ελληνικής γλώσσας*, επιμ. Α.-Φ. Χριστίδης *et al.* Αθήνα: ΥΠΕΠΘ & Κέντρο Ελληνικής Γλώσσας, 39-46.
- [11] Longobardi G. (in print). Greek, Italian and their Dialects: a Quantitative Taxonomy. *Πρακτικά του 3ου Διεθνές Συνέδριου Νεοελληνικών Διαλέκτων και Γλωσσολογικής Θεωρίας* (Λευκωσία, Κύπρος, 2007).
- [12] Morosi G. (1970). *Studi sui dialetti greci della Terra d'Otranto*, Lecce: Tip. Ed. Salentina.
- [13] Parlangei P. (2007). Note per la storia del griko. *Studi Linguistici Salentini* 31: 5-89.
- [14] Προφίλη Ό. (1986). Description du système phonétique et phonologique du parler grico de Corigliano d'Otranto. *Studi Linguistici Salentini* 14.
- [15] Προφίλη Ό. (1999). Η Ελληνική στη Νότια Ιταλία. *Διαλεκτικοί θύλακοι της Ελληνικής Γλώσσας*, επιμ. Α.-Φ. Χριστίδης *et al.*, Αθήνα-Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας, 31-38.
- [16] Rohlfs G. (1972). Persistenza della grecità nell'Italia meridionale. *Studi e Ricerche su Lingua e Dialetti d'Italia*. Firenze: Sansoni, 231-245.
- [17] Romano A. (2000). Convergence and divergence of prosodic subsystems of the dialects spoken in the Salento (Italy) - a linguistic and instrumental approach. *Proceedings of the "1st ICLaVE"* (Barcelona, Ισπανία, 2000), 168-178.
- [18] Romano A. (2004). Uno spoglio fonetico della base di dati-audio « The Græcanic Lexicon » dell'Università di Patraso. In A. De Dominicis, L. Mori & M. Stefani (eds.) *Proceedings of the "XIV Giornate del G.F.S."* (Viterbo, Italy 2003), Roma: Esagrafica, 81-86.
- [19] Romano A. (in print). Acoustic data about the Griko vowel system, *Πρακτικά του 3ου Διεθνές Συνέδριου Νεοελληνικών Διαλέκτων και Γλωσσολογικής Θεωρίας* (Λευκωσία, Κύπρος, 2007).
- [20] Romano A. (2008). *Inventari sonori delle lingue: elementi descrittivi di sistemi e processi di variazione segmentali e sovrasegmentali*, Alessandria: Dell'Orso.
- [21] Romano A. & Marra P. (2008). *Il griko nel terzo millennio: « speculazioni » su una lingua in agonia*, Parabita (Lecce): Il laboratorio.

- [22] Sobrero A. & Miglietta A. (2007). Spie morfologiche della resistenza del grico nel Salento, *Rivista Italiana di Dialettologia*, XXXI: 19-28.
[23] Tommasi S. (1996). *Katalisti o kosmo. Materiali per un seminario sul griko, Ghetonìa-Calimera*, Galatina: Ed. Salentina.

Προσάρτημα

Ενότητες διαλόγων από "Το vikènde" (βλέπε [23]): κείμενο σε σημειώσεις Γκρίκο και Ιταλικά

- 01_01a Deste tue,
Guarda queste!
- 01_01b ti steu kkaledde!
Che sono carine!
- 01_01c Kaisato na termànete sto kamatzulài?
Vi siete sedute a scaldarvi al solicello?
- 01_01d Guikato ston ijo **kundu** e stavvrikule!
Siete uscite al sole come le lucertole!
- 02_01 Mi ppronì frontì tu marti guenni t'afidi pu kau sto lisari.
Al primo tuono di marzo esce la serpe da sotto la pietra.
- 03_02a Achà, echi **puru** o tio Ciseppo ettù!
Ahà, c'è pure lo zio Ciseppo qui!
- 03_02b Makarriòttu:
Beato lui:
- 03_02c fènete ka e chroni tunù 'e ttu diaennu' mmakà pu panu.
sembra che gli anni di lui non gli passino affatto di sopra.
- 04_01a Dela, Peppina.
Vieni, Peppina.
- 04_01b Sìmmèri ettumpì e' ssa **pparadiso**. Stèome alion òrrie sto termuddhi.
Oggi qui è un paradiso. Stiamo un po' belle al calduccio.
- 05_01a O **justo**,
Il giusto,
- 05_01b **dopu** i tzichra, es frontè ce ta nnerà pu e' janomena ittes addomè,
dopo il freddo, i tuoni e le acque che sono maturate le scorse settimane,
- 05_01c t'ùsele **pròbbio** lion ijo na stannòsome.
ci voleva proprio un po' di sole per farci asciugare.
- 06_01 Valosti na mas kordosi nerò, e **vekkia**, fetò.
Si è messa a riempirci d'acqua, la vecchia, quest'anno.
- 07_02a Prai kascio, Peppina: mi mmini **mpalàin** ecirtèa.
Vieni a sederti, Peppina: non stare impalata di là.
- 07_02b Pu ènna pai panta pratonta?
Dov'è che vai sempre girando?
- 07_02c Pame na su **mpoggetzi** lion o poda.
Andiamo che ti "appoggi" un po' i piedi.
- 08_02a Ka fidete na stasi kammèni arte e Peppina?
Che riesce a stare seduta adesso la Peppina?
(cp. Sal. ca se fide cu sse stescia ssettata mo' la Peppina?)
- 08_02b Pu na **skiattetzi**!
Che schiatti!
- 08_02c Manechò o tànato teleste ka in **empoggei**, cini.
Solo la morte può farla "appoggiare", quella.
- 09_02 Eh, Ninamu, ka panta **vòtise** àscima os kristianò!
Eh, Nina mia, ca sempre ti giri (rivolgi) male ai "cristiani"!
- 10_03 Mìnone, Peppina, arte ka s'ida: ènna se rotiso tz'ena pprama. Esù **certo** ka tzeri.
Aspetta, Peppina, adesso che ti ho vista: ho da chiederti di una cosa. Tu certo la sai.
- 11_03a Tì ènna tzero, Mundanzia?
Che ho da sapere, Mundanzia?
- 11_03b Pemma, ettù steo.
Dimmi, qua sto.

- 12_04 Deje, dela ttusumuddhia. 'E ttelo na mas kusi tispo.
No, vieni qua vicino. Non voglio che senta nessuno.
- 13_04 Vuvvù! E' pprama krifò?
Vuvvù! È cosa segreta?
- 14_05a Deje. 'E ttelo na mas echu' tti pi.
No. Non voglio che ci abbiano da dirla.
- 14_05b Pa' ce pistèune ka **piànnome** a **penserria** os addhò.
Vanno e credono che prendiamo i pensieri degli altri.
- 15_03 Mme! Ce 'e ssòzzome pi cippu tèlome arte?
Beh! E non possiamo dire quello che vogliamo adesso?
- 16_06a Kuse: attepornà,
Ascolta: stamattina,
- 16_06b ida i kkiatera i Mmakolata mo cciùritti.
vidi la figlia della Mmacolata con suo padre.
- 16_06c Pos ene?
Com'è?
- 16_06d Ìtzera ca iche fionta.
Sapevo che era "scappata".
- 16_06e Poan ione? Mino. I pprassei pu ddiake, mu fènete.
Quando fu? Aspetta. Il venerdì che passò, mi pare.
- 17_05a Fionta?
"Scappata"?
- 17_05b Deje, ti lei?!
No, che dici?!
- 17_05c Vriskese **frastornài**.
Ti trovi frastornata.
- 17_05d Ka stèun oli mia!
Che stanno tutti una!
- 18_07a1 Umme, evo t'ùkusa. M'o t'ùpe mia kristiani ka stei poddhì **kùkkia**.
Sì, io l'ho sentito. Me l'ha detto una cristiana che sta molto vicina.
- 18_07a2 **Pròbbio** in addomà pu ddiake.
Proprio la settimana che passò.
- 18_07b M'ùpane ka tàrasse sto **largo**, 'en etzero pu, mon **annamurào**.
Mi hanno detto che parti al largo, non so dove, con l'innamorato.
- 19_06a Tuo en jalìssio, Mundantzia.
Questo è vero, Mundantzia.
- 19_06b **Ma** de' kka fiane.
Ma non che fuggirono.
- 19_06c Pirtane na kamu' to **vikende**.
Sono andati a fare il week-end.
- 20_08 Pose, pose? **Vikende**? Ti praman e' ttuo mapale?
Come, come? Week-end? Che cosa è questa di nuovo?
- ...
- 38_11a An enghia tuo, Peppina...
Se è questo, Peppina...
- 38_11b **Certo**, emì **sordu** 'en ìchamo, ka toa pina iche.
Certo, noi soldi non avevamo, che allora fame c'era
- 38_11c Ce e mane, **addho ka ssordu**: mas kratènnane toso ta mmaddia panu.
E le madri, altro che soldi: ci tenevano tanto d'occhi sopra.
- 38_11d **Però**, a ttelùsamo... An e pròbbio telùsamo...
Però se volevamo... se proprio volevamo...
- 38_11e Pos èna su po? A ttelùsamo, **puru** emì o kànnamo, to **vikende**.
Che ti devo dire? Se volevamo, pure noi lo facevamo, il week-end.

Object clitics in the Modern Greek dialects of Asia Minor: diachronic and dialectal variation in the encoding of argument structure

ANNETTE TEFFETELLER
Concordia University
Montreal, Quebec, Canada
teffet@alcor.concordia.ca

1. Introduction

The pronominal object clitics of Greek, Ancient and Modern, have long been of interest to syntacticians and historians of the language.⁷¹ Janse (2008) provides a useful collection of data but the theoretical dichotomy which he reports (Janse 2008:166; for discussion of his view of clitics as ‘a category *sui generis*’ see Janse 1998a, 1998b) has in fact been superseded. It is no longer necessary to choose between the two positions exemplified for him on the one hand by Irene Philippaki-Warburton (1977, 1987; Philippaki-Warburton and Spyropoulos 1999; Philippaki-Warburton et al 2004) and on the other by Brian Joseph (2001, 2002a, 2002b, 2002c) with regard to Modern Greek, namely, that a) clitics are words and themselves arguments of the verb, with their associated NPs categorized as adjuncts (the Philippaki-Warburton position), or that b) clitics are affixes and consequently mere agreement markers, with the argument category filled by the associated NP (Joseph’s position). Kenneth Hale (2003) and Marianne Mithun (2003), independently, have shown, with data from Navajo and Yup’ik, that pronominal *affixes* can themselves function as core arguments of the verb. Janse (2008:166) observes that ‘the status of clitic pronouns is not the same for all the dialects and stages of the Greek language’. It is indeed an observational fact that in the history of Greek the placement and the obligatoriness of object clitics has varied but whether the actual function of clitic pronouns has changed is still a matter for investigation. It is here suggested that the diachronic and dialectal variations in clitic placement and obligatoriness from Ancient to Modern Greek have no bearing on the *status* of clitic pronouns, that is, the function of object clitics in the argument structure of the language. The investigation will start with subject reference in the Indo-European verbal morphology.

2. Subject reference

Notwithstanding its well-known genetic anomalies, Anatolian of the second millennium BC shows the typical Indo-European distribution of double marking for subjects of a clause, both on the verb and optionally on independent nominal or pronominal items, e.g.:

- (1)
a. *ta=âmaâ œurtiya[llan par]Ω Ēpmi DUMU.É.GAL âuppi wΩtar parΩ epzi [LUGAL]-i*
SAL.LUGAL=ya StBoT 8, I 13-15
‘I hold out to them a dish, a palace attendant holds out pure water, to the king and queen.’

⁷¹ Recent studies include Alexiadou & Anagnostopoulou 2000, Anagnostopoulou 1999, Condoravdi & Kiparsky 2001, Janse 1994, 1998a, 1998b, 2008, Joseph 1988, Kallulli 2000, Pappas 2004, Philippaki-Warburton 1977, Philippaki-Warburton & Spyropoulos 1999, Philippaki-Warburton et al. 2004, Revithiadou & Spyropoulos 2008, Taylor 1996, cf. Holton, Mackridge & Philippaki-Warburton 1997, Siewierska 1999, Zwicky 1977; on Asia Minor Greek see Janse 1994, 1998a, 1998b, 2002, 2004, 2008; and see further references in the following text.

b. „**k=wa** LUGAL-*uâ=âmiâ kiâæa*
'I will become your king'

c. **zik=wa** UR.BAR.RA-*aâ kiâtat*
'you have become a wolf.'

Hittite is by far the most copiously attested of the Anatolian languages (and consequently the best understood) and like other early Indo-European languages (and some modern ones such as Spanish), Hittite has been classified within the framework of mainstream generative syntactic theory as a null subject language, with the person markers on the verb interpreted as agreement morphology co-referencing an independent subject which functions as the argument of the verb and which may be overt (a lexical nominal or an independent pronoun) or, optionally, null. In other frameworks, outside the mainstream of current syntactic theory, verbal subject markers of the Indo-European type are interpreted as incorporated pronouns (originating in independent pronouns first postcliticized to the verb and eventually fully incorporated) and functioning as core arguments of the verb. (IE verbal markers originate in pronominal forms: Szemerényi 1996, Bomhard 1988, Sihler 1995; for pronominal affixes as arguments see Hale 2003 and Mithun 2003 – which implicitly answers the objections of Bresnan & Mchombo 1987 to Jelinek 1984; cf. Evans 1999; for Hittite see Hoffner and Melchert 2008.)

Since the Indo-European verb marks only subject reference, direct objects in Hittite (and across Indo-European), appearing optionally as lexical nominals or as full or clitic pronouns (pragmatically conditioned), the latter view would categorize Hittite (and Indo-European generally) as a ‘mixed’ type of language in the sense of Jelinek (1987).

English in mainstream syntactic theory is considered to have an obligatorily overt subject with finite verbs (with the exception of the imperative), although colloquial English abounds with ‘subject-less’ clauses, e.g.: ‘Beats me’, ‘Don’t know’, ‘Told you so’, ‘Been there, done that’, ‘Ran’, ‘Finished?’, ‘Cheats’. The circumstance in which a language such as English that admits verbs without an overt subject is considered to be *obligatorily* marked with an overt subject while at the same time a language such as Hittite that *never* admits finite verbs that are *not* overtly marked for subject is considered to be a null subject language might, one would think, prompt theorists of this persuasion to reconsider their theoretical categories with regard to subject marking. And object marking as well, which we will return to, but first some data from Greek.

Ancient Greek – from the Core Indo-European group (excluding Anatolian and Tocharian) – Ancient Greek of all periods, from Mycenaean to the Koine, admits structures which have (or appear to have) a lexical nominal or an independent pronoun in the subject position, in addition to the person marking in the verbal morphology. Here is an example with an embarrassment of riches in the ‘subject’ category; we will see it again later:

(2)
 αὐτὰρ ὁ βοῦν ἱέρευσευ ἄναξ ἀνδρῶν Ἀγαμέμνων II. 2.402
 but he ox he-sacrificed lord of men Agamemnon.

But Greek of all periods also permits a finite verb standing alone as a clause, with no overt, independent ‘subject’ designated by a nominal or pronominal separate from the verb, e.g., *eijmiv, fhsiv*; cf. Latin *venio, cogito*. Hittite likewise, e.g.: *Ēâmi* ‘I am’, *aršanieš* ‘you were envious’, *memiškeš* ‘you were saying’ *šipanti* ‘he libates, offers’, *waâtai* ‘he sins’, *adanzi* ‘they eat’.

As the great French Indo-Europeanist Antoine Meillet observed well over a century ago, the syntactic system of the archaic Indo-European languages is based on the principle that he articulated as “the autonomy of the word” — a function of semantics and of the inflectional morphology:

La phrase indo-européenne se composait de mots autonomes, dont chacun suffisait à exprimer un sens complet et la fonction remplie. ... Outre le sens exprimé par le thème, la flexion marque le rôle joué par chaque mot dans la phrase; le mot est donc autonome et suffit par lui-même à indiquer son sens et son rôle dans le discours (Meillet 1903 [1937]:439, 356)

Apart from the verbal morphology, Core Indo-European had only nominals or full pronominals as independent 'subjects' (that is, 'subjects' on a traditional analysis – and on a mainstream analysis today); it had no pronominal subject clitics. Here Anatolian has innovated in the creation of a third-person definite referential clitic pronoun, marked for gender, common and neuter, but restricted to a particular class of (predominantly stative) intransitive verbs, the so-called 'unaccusatives' (Garrett 1990a, 1990b, 1996). This clitic pronoun is in complementary distribution with both the full (emphatic) demonstrative pronoun used for third-person reference

and lexical 'subjects'; and while 'clitic doubling' has been claimed for Anatolian, e.g. Luwian (Melchert 2003:201):

(3) *tŌñ-ti(y)-ata malli aiyaru tapŌruwa hŕr,ta tatarriyamna*
'Let them, the *t*, oaths and curses, become oil (and) honey' (KUB 9.6+ ii 12-13),

support is growing for the view that 'right-dislocation' and 'clitic doubling' are not in fact to be seen in such structures (Melchert 2010:2, with reference to Bauer forthcoming): 'right dislocated NP appears to be mere apposition to anaphoric pronoun. Thus no true "right dislocation" separate from extraposition, merely extraposition of epexegetic NP that is apposition to pronoun, *which is the real argument*.'

On the basis of the evidence we have, it would seem preferable – indeed inescapable – to classify Indo-European with respect to the 'subject' (or primary argument – and not separating out here 'agent' from 'subject') as (in generative terms) 'head-marking' as opposed to 'dependent-marking' (terms I prefer to avoid since they entail the full theoretical model, but will use for the moment), with optional lexical or pronominal adjuncts (or co-referents) to the subject reference that is fully marked in the verbal morphology. Indo-European would then be classed (by those who accept this type) as a pronominal argument language, at least insofar as the 'subject' is concerned.

3. Direct object reference

Given that IE marks subject reference on every finite verb, what then of the direct object? Of course many languages that mark subject reference on the verb or in an obligatory (second-position) clitic string mark object reference as well (often indirect as well as direct), e.g.:

(4) Mohawk:
leksa:'a raksa:'a wahonwa:ienhte'
girl boy she-hit-him
'The girl hit the boy.'

leksa:'a raksa:'a wahshako:ienhte'
girl boy he-hit-her
'The boy hit the girl.'

(5) Classical Arabic (Lambrecht 2001):
[Halidun], qabaltuhu l-yawma
Halid.NOM met.1SG.3SG.ACC the-day.ACC
'Halid, I met him TODAY.'

(6) Navajo (Hale 2003: see discussion below):

ni-sh-hozh.
2SG-1SG-tickle
'I tickle you.'

It is worth quoting at length from Kenneth Hale's article 'On the significance of Eloise Jelinek's Pronominal Argument Hypothesis' (Hale 2003:12-13;cf. Jelinek 1984, 1987, 2006):

In a language belonging to the PA type, the person-number morphology internal to a verb word represents the direct arguments of the verb. These elements are not agreement morphology. Instead they *are* the arguments, pure and simple. In the Navajo verb word (or rather, somewhat more accurately, "verb sentence") cited in (1) below, the prefixes *ni-* and *sh-* are, respectively the object and subject of the clause:

(1) *ni-sh-hozh.*
2SG-1sg-tickle
"I tickle you."

There are no "small pro" elements in this sentence, and if an independent pronoun appeared, as in (2), it would not be an argument but rather it would be a contrastive adjunct:

(2) *Ni ni-sh-hozh.*
"I tickle YOU."

The independent pronoun *ni* 'you' is, to be sure, *linked* to the verb-internal object *ni-*, but it is not an argument of the verb, any more than the first *you* is an argument of *tickle* in the English *as for* construction in (3):

(3) As for you, I'm tickling you.

In short, the Navajo independent pronoun *ni* 'you' in (2) is not related to the prefix *ni-* in the way an argument is related to agreement morphology. It is the prefix, not the independent pronoun, that represents — alone and fully — the object argument of the verb. The same can be said of a nominal expression, like 'awéé' 'baby' in (4):

(4) 'Awéé' bi-'nii-sh-hóósh.
baby 3INCH-1SG-tickle
"I start to tickle the baby."

This is an inchoative verb form, with the direct object of the verb appearing as *bi-* directly before the inchoative morphology *-'nii-* (glossed INCH)-. Here again, the true arguments are represented by the verb-internal person-number morphology, i.e., the third person object pronoun *bi-* and the first person singular subject pronoun *sh-*. The nominal 'awéé' 'baby' is an adjunct, not an argument of the verb. Its structural relation to the sentence can be compared to that of the English left-dislocated nominal *the baby* in (5), where the true object argument is the resumptive pronoun *him*:

(5) The baby, I will start to tickle him/her.

The idea, then, is that Navajo is a language in which all of the arguments of a verb are pronouns and, further, the pronouns in question are morphologically dependent (i.e., they are affixes, inflection). The verb word is in reality a complete sentence — a "verb sentence" (VS), although the more conventional (albeit less accurate) expression "verb word" will be occasionally used throughout this discussion.'

Hale's analysis is echoed by Mithuen (2003):

In Navajo as in Yup'ik, obligatory pronominal affixes on every verb identify the core arguments of the clause. (258)

...in languages with pronominal affixes, each verb constitutes a complete minimal clause in itself, the skeleton or nucleus of the clause. (274)

Agreement is of course not the primary function of pronominal affixes; their role is to evoke referents. (276)

Compare Abkhaz on the one hand and French on the other:

(7) Lambrecht 2001:(49)

Abkhaz

[a-xàc'a]_i [a-ph^oðs]_j

ART-man ART-woman

[a-s^oq^o'ð]_k ø_k-lð_j-y_i-te-yt'

ART-book it-to.her-he-gave-TNS

'The man gave the book to the woman.'

(Lambrecht 2001:1061)

Lambrecht notes that 'it is revealing to compare (49) with its (slightly modified) spoken French equivalent in (49)'; to emphasize the formal similarity with (49), the morphologically bound status of the pronouns in the French verb complex is indicated by hyphens:

(8) Lambrecht 2001:(49')

[L'homme]_i [cette femme-là]_j

the man that woman-there

[mon livre]_k il_i-le_klui_j-a-donné

my book he-it-to.her-has-given

'The man he GAVE my book to that woman.'

Though perhaps pragmatically unusual, (49') is nevertheless a grammatically well-formed spoken French sentence. The striking structural similarity between (49) and (49') confirms the observation, made early on by Vendryès (1914), that modern spoken French is typologically close to certain polysynthetic languages (Vendryès compares French to the Amerindian language Chinook) (Lambrecht 2001: 1061).

Of course in Navajo and Yup'ik, as in Abkhaz (and evidently in French), the object is morphologically marked on the verb, whereas early Indo-European uses for direct object reference (whether core or adjunct is the issue here) lexical nominals and independent pronouns, which may be separated from the verb, as well as clitic pronouns and null instantiation of pronominals.

Ancient Greek has optional instantiation of object reference in a clitic pronoun, on occasion combined with extraposed lexical reference in adjunct position, e.g.:

(9)

δὴ γὰρ μιν ἔφαντ' ἐπιδήμιον εἶναι,

indeed for him they-said among-people to-be,

so;n patevr j

your father

'For indeed they said that he was among his people, /

your father'

Homer *Odyssey* 1.194-195

This is a structure seen already in Mycenaean Greek in the second millennium BC:

(10)

PY Ep 704.5

δα-μο-δε-μι πα-σι • ο-να-το ε-κε-ε

δ³μοσ δε μιν π¹σι • ον²τον εκ¹εην

'but the δ³μοσ says that she holds the lease'

(cf. Ruijgh 1967:30, Ventris & Chadwick 1973: 254, Janse 2008:173 mistakes the syntactic structure: the infinitive *ek^hehen* depends on the object clitic pronoun *min*, not the other way around.)

The Mycenaean clitic is in Wackernagel's position, forming a 'word' with *dōmos* and *de*.

The second-millennium Anatolian languages have lengthy sentence-initial 'chains' of enclitic particles and anaphoric pronouns, with up to six places in Hittite:

- 1) connectives (e.g. adversative =*ma*, adds new information)
- 2) quotative particle (=wa(*r*))
- 3) 3rd person object (of transitive, e.g. =*an*) or subject (of 'unaccusative' intransitive)
- 4) 3rd person dative, 1st/2nd person (e.g. =*āi*)
- 5) reflexive particle (=za)
- 6) local (/aspectual) particles (e.g. =*kan*)

The Luwian clitic chain shows a slight variation from the Hittite but is equally fixed:

- 1) conjunction -*ha* or -*pa*
- 2) quotative particle -*wa*-,
- 3) dative or reflexive pronoun
- 4) nominative or accusative pronoun
- 5) local particle

The Greek clitic pronoun *min* has anaphoric reference and is sometimes associated with a lexical NP in adjunct position either to its left or to its right. In one case it is used with *αὐτόν*~ with reflexive force:

(11)

Od. 4.244

αὐτόν μιν πληγῆσιν ἀεικελίησι δαμάσσᾱ

Helen of Odysseus: 'disfiguring himself with grievous blows'

(12)

Anaphoric reference:

a. *Od.* 10.210-12

εὔρον δ' ἐν βήσσησι τετυγμένα δώματα Κίρκῃ

ξεστοῖσιν λάεσσι, περισκέπτω ἐνὶ χώρῳ·

ἀμφὶ δέ μιν λύκοι ἦσαν ὀρέστεροι ἢ δὲ λέοντῆ

'They found in the glades the built halls of Circe,

with polished stones, in an open clearing;

and around it there were mountain wolves and lions.'

(anaphoric reference to the neuter plural *dōvmata* 'halls' in the previous clause)

b. *Od.* 1.194-5

νῦν δ' ἦλθον· δὴ γάρ μιν ἔφαντ' ἐπιδήμιον εἶαι,

σὸν πατέρ'

Athena in disguise as Mentès to Telemachus, who has asked about Mentès' relation to his father, Odysseus; Athena/Mentès replies with a statement of his identity and his relationship to Odysseus and continues: 'And now I have come, for they said that he was among his people, your father.'

(13)

Il. 1.100

τότε κέν μιν ἰλασσάμενοι πεπίθοιμεν.

Calchas to the Achaeans regarding Apollo's anger: the cause is Agamenmon's dishonouring the priest Chryses and refusing to accept the ransom for his daughter; the god will not relent until we return the girl freely and sacrifice a hecatomb to him at Chryse: 'then, appeasing him, perhaps we might persuade him.'

(anaphoric reference of the clitic pronoun, referring to e]khbovlo~ in line 96 (and all the subsequent anaphoric references encoded in the ensuing verbal morphology))

In addition to independent and clitic pronouns, Greek also allows null instantiation of object pronouns, as do many languages both within and outside the Indo-European family; for Navajo see Mithun (2003:258). The following examples illustrate the situation in Homeric and Classical Greek, in poetry and prose:

(14)

a. Φοίνικε δ' ἄγον ἄνδρῃ.

'Phoenician men brought [it]. '

Homer *Iliad* 23.744

b. οὐκ ἐκεῖνδ' ἄλλ' ἐκείνη κείνον ἐνθάδ' ἤγαγεν.

'He did not bring [her] here, but she brought him.'

Euripides *Orestes*

742

c. ὁ δὲ ἐμπιμπλᾷ ἀπάντων τὴν γνώμην ἀπέπεμπε.

'Having satisfied the minds of all he dismissed [them].'

Xenophon *Anabasis* 1.7.8

d. ἀράττω.

'I assail [him].'

Aristophanes *Clouds* 1373

In languages outside the Indo-European family, Turkish, among others, shows null instantiation of object pronouns, e.g.:

(15)

a. *Adam tas^ -i oglan-a at-ti*

man stone-ACC boy-DAT throw-PAST

'The man threw the STONE at the BOY.'

b. *Adam ~oglan-a at-ti [tas^-i]i*

man boy-DAT throw-PAST stone-ACC

'The man threw it at the BOY, THE STONE.'

c. *Adam tas^-i ~at-ti [oglan-a]i*

man stone-ACC throw-PAST boy-DAT

'The man threw the STONE at the BOY.'

('The man threw the STONE at him, the BOY.)

(Lambrecht 2001:1056).

A question arises at this point: when a language permits null instantiation of objects, the semantics of the verb nevertheless make it clear that an object is to be understood and typically the discourse context makes it more or less clear what that object is – ideally, inescapably clear, but in every instance at least reasonably clear. If an object is to be understood from the semantics of the verb (one doesn't just *hit*, one hits *something*), and if, moreover, what that object is is on the whole clear from the discourse context, then we must consider whether the null phonological representation is to be understood as the default or core structure. If so, the overt expression of the object, whether (in ascending order of emphasis and specification) by a clitic pronominal, a full independent pronoun, or a lexical nominal, would in every case constitute an adjunct to the null pronominal

reference. The object, being inherent in the verb, would be viewed as marked on the verb, in a null phonological representation. On such a model Indo-European would be classed as a full pronominal argument language.

We are then brought back to our earlier example, from Homeric Greek:

(16) αυjta;r oj bou`n ijevreusen a[nax ajndrw`n jAgamevmnwn Il.
2.402

where the nuclear clause is now seen to be ijevreusen ‘he-sacrificed-it’, with all lexical items, and the pronoun oj as well, in adjunct position.

4. Asia Minor Greek

And we arrive at last at the Greek of Asia Minor. R.M. Dawkins, who recorded the dialects *in situ* before the ‘exchange of populations’ in the 1920s, gives us our most reliable view of the Asia Minor dialects before their natural development was disrupted and their speakers dispersed (Dawkins 1916, 1931, 1937, 1940; for Pontic cf. Drettas 1997). He is particularly eloquent on the trajectory to be seen in the position of pronominal objects across the Greek dialects:

The fact seems to be that the position of the pronominal object forms a chain right across the Greek world. In Italy and on the mainland the object always precedes; in Crete and all the islands as far as Cyprus it may follow, but only in positive main clauses; in Cappadocia it must follow the verb in positive, but never in negative or dependent clauses; at Pharasa in the Taurus the object follows even in negative sentences, and lastly and finally in Pontos it always follows even in dependent clauses and one finds for example that *I want to say it* appears as qevlw na; levgw to, a word order absolutely unheard of and impossible anywhere else in the whole Greek world’ (Dawkins 1940:22-23).

(17) Pontic Greek

a. *egó séna dilévo=se*

I you I-feed=you

‘I will feed you’

(Dawkins 1916:314; cf. Janse 1998:538)

The ‘quasi-Pontic dialect’ of Pharasa (Dawkins 1940:5)

b. *to=mávro dekanínken=da to=pózi*

the=grey he-bit=it the=black

‘the grey one was biting the black one’

(Dawkins 1916:558; cf. Janse 1998:540)

Pontic had by this point developed obligatory object markers on the verb to parallel the subject markers Greek inherited millennia earlier from Indo-European. This is a structure that was optional at all periods of Greek (with variation in placement) and is widely used in Standard Modern Greek but it was evidently made obligatory in all contexts only in Pontic and the closely related dialects of Asia Minor Greek.

Devine and Stephens (2000:158) note that above all the schema Alcmánicum attests to the status of early Greek as a pronominal argument language, e.g.:

(18) ἦχι ροᾶ Σιμόεϊ συμβάλλετον ἠδὲ Σκάμανδρῶ Il. 5.774
where their streams Simoeis they-two-merge and Skamandros
‘where the Simoeis and the Scamander merge their streams’.

Greek developed configurational syntactic structures already in the ancient period. But in its argument structure it has evidently retained its early typology. And the modern Greek dialects of eastern Asia Minor, Pontic above all, simply made overt the argument structure inherent in the language from its earliest attestation.

References

- Alexiadou A. & Anagnostopoulou E. (2000). Clitic-doubling and non-configurationality, *Proceedings of the North East Linguistic Society* 30, vol. 1: 17-28. Amherst, MA.
- Anagnostopoulou E. (1999). Conditions on Clitic Doubling in Greek. In H. van Riemsdijk (ed.) *Clitics in the Languages of Europe*, 761-798. Berlin.
- Bauer A. (forthcoming). Verberststellung im Hethitischen. In T. Krisch and T. Lindner (eds.) *Acts of the XIIIth International Conference of the Society of Indo-European Studies, Salzburg, September 21st to September 27th, 2008*.
- Bomhard A. R. (1988). The Prehistoric Development of the Athematic Verbal Endings in Proto-Indo-European. In Y. Arbeitman (ed.) *A Linguistic Happening in Memory of Ben Schwartz*: 475-88. Louvain-la-Neuve.
- Bresnan J. & Mchombo S. A. (1987). Topic, Pronoun, and Agreement in Chichewa, *Language* 63: 741-782.
- Condoravdi C. & Kiparsky P. (2001). Clitics and clause structure. *Journal of Greek Linguistics* 2: 1-39.
- Dawkins R.M. (1916). *Modern Greek in Asia Minor: A Study of the Dialects of Silli, Cappadocia and Phárasa with Grammar, Texts, Translations and Glossary*. Cambridge.
- Dawkins R.M. (1931). Notes on the Study of the Modern Greek of Pontos. *Transactions of the Philological Society* 31: 389-400.
- Dawkins R.M. (1937). The Pontic Dialect of Modern Greek in Asia Minor and Russia. *Transactions of the Philological Society* 36: 15-52.
- Dawkins R.M. (1940). The Dialects of Modern Greek. *Transactions of the Philological Society* 39: 1-38.
- Devine A.M. & Stephens L.D. (2000). *Discontinuous Syntax: Hyperbaton in Greek*. New York and Oxford.
- Drettas G. (1997). *Aspects Pontiques*. Paris.
- Evans N. (1999). Why argument affixes in polysynthetic languages are not pronouns: evidence from Bininj Gun-wok, *Sprachtypologie und Universalienforschung* 52: 255-281.
- Garrett A. (1990a). *The Syntax of Anatolian Pronominal Clitics*. Doctoral dissertation, Harvard University.
- Garrett A. (1990b). 'Hittite Enclitic Subjects and Transitive Verbs', *Journal of Cuneiform Studies* 42: 227-42.
- Garrett A. (1990c). The Origin of NP Split Ergativity, *Language* 66: 261-296.
- Garrett A. (1996). Wackernagel's Law and Unaccusativity in Hittite. In A. L. Halpern and A. M. Zwicky (eds.) *Approaching Second: Second Position Clitics and Related Phenomena*: 85-133. Stanford.
- Hale K. (2003). On the Significance of Eloise Jelinek's Pronominal Argument Hypothesis. In A. Carnie, H. Harley, and M. Willie (eds.) *Formal Approaches to Function in Grammar: In Honor of Eloise Jelinek*: 11-44. Amsterdam.
- Hoffner H. & Melchert H.C. (2008). *A Grammar of the Hittite Language*. Winona Lake, IN.
- Holton, D., P. Mackridge & I. Philippaki-Warbuton. 1997. *Greek: A Comprehensive Grammar of the Modern Language*. London.
- Janse M. (1994). Son of Wackernagel. The distribution of object clitic pronouns in Cappadocian. In I. Philippaki-Warbuton, K. Nicolaidis, and M. Sifianou (eds.) *Themes in Greek Linguistics* Vol. 1: 435-442. Amsterdam.
- Janse M. (1998a). Cappadocian Clitics and the Syntax-Morphology Interface. In B.D. Joseph, G.C. Horrocks & I. Philippaki-Warbuton (eds.) *Themes in Greek Linguistics* Vol. 2, ed., 257-281. Amsterdam.
- Janse M. (1998b). Grammaticalization and Typological Change: The Clitic Cline in Inner Asia Minor Greek. In M. Janse (ed.) *Productivity and Creativity: Studies in General and Descriptive Linguistics in Honor of E. M. Uhlenbeck*, 521-547. Berlin.
- Janse M. (2002). Aspects of Pontic grammar. *Journal of Greek Linguistics* 3: 203-231.
- Janse M. (2004). Animacy, Definiteness and Case in Cappadocian and Other Asia Minor Greek Dialects. *Journal of Greek Linguistics* 5: 3-26.

- Janse M. (2008). Clitic Doubling from Ancient to Asia Minor Greek. In D. Kallulli and L. Tasmowski (eds.) *Clitic Doubling in the Balkan Languages*, 165-202. Amsterdam.
- Jelinek E. (1984). Empty Categories, Case, and Configurationality, *Natural Language and Linguistic Theory* 2:39-76.
- Jelinek E. (1987). Auxiliaries and Ergative Splits: A Typological Parameter. In M. Harris and P. Ramat (eds.) *Historical Development of Auxiliaries*, 85-108. Berlin.
- Jelinek E. (2006). The Pronominal Argument Parameter, in *Arguments and Agreement*, ed. P. Ackema et al, 261-88. Oxford.
- Joseph B. D. (1988). Pronominal Affixes in Modern Greek: The Case against Clisis,. In Diane Brentari, Gary Larson, and Lynn A. MacLeod (eds.) *Papers from the 24th Regional Meeting, Chicago Linguistic Society*, ed., 203-215. Chicago.
- Joseph B. D. (2001). Dialect Evidence Bearing on the Definition of 'Word' in Greek. In *Proceedings of the First International Conference of Modern Greek Dialects and Linguistic Theory*, ed. by A. Ralli, B.D. Joseph, and M. Janse, 89-104. Patras.
- Joseph B. D. (2002a). Defining 'Word' in Modern Greek. A Response to Philippaki-Warburton and Spyropoulos 1999. *Yearbook of Morphology 2001*: 87-114.
- Joseph B. D. (2002b). On Defining 'Word' in Modern Greek. In C. Clairis (ed.) *Recherches en linguistique grecque*. Vol. 1: 247-250. Paris.
- Joseph B. D. (2002c). 'Word' in Modern Greek, in *Word*. In R.M.W. Dixon and A. Aikhenvald (eds.) *A Cross-Linguistic Typology*, 243-265. Cambridge.
- Kallulli D. (2000). Direct Object Clitic Doubling in Albanian and Greek. In F. Beukema and M. den Dikken (eds.) *Clitic Phenomena in European Languages*, 209-248. Amsterdam.
- Lambrecht K. (2001). Dislocation. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds.) *Language Typology and Language Universals, Handbücher zur Sprach- und Kommunikationswissenschaft*, 20, Vol. 2: 1050-1078. Berlin.
- Meillet A. 1903 (1937). *Introduction à l'étude comparative des langues indo-européennes*, 8th ed. 1937. Paris.
- Melchert H. C. (2003). Language. In H.C. Melchert (ed.) *The Luwians*, 170-210. Leiden.
- Melchert H. C. (2010). Syntax and Prosody in Hittite Word Order. *Language Variation & Change Workshop*, University of Chicago, January 15, 2010.
- Mithun M. (2003). Pronouns and Agreement: The Information Status of Pronominal Affixes, *Transactions of the Philological Society* 101: 235-78.
- Pappas P. (2004). *Variation and Morphosyntactic Change in Greek. From Clitics to Affixes*. Basington.
- Philippaki-Warburton I. (1977). Modern Greek Clitic Pronouns and the 'Surface Syntax Constraints' Hypothesis. *Journal of Linguistics* 13: 259-281.
- Philippaki-Warburton I. (1987). The Theory of Empty Categories and the Pro-drop Parameter in Modern Greek. *Journal of Linguistics* 23: 289-318.
- Philippaki-Warburton I. & Spyropoulos V. (1999). On the Boundaries of Inflection and Syntax: Greek Pronominal Clitics and Particles. *Yearbook of Morphology 1998*, Vol. 8, ed. G. Booij and J. Van Marle, 45-72. Dordrecht.
- Philippaki-Warburton, I., S. Varlokosta, M. Georgiafentis, and G. Kotzoglou. 2004. Moving from Theta-positions. Pronominal Clitic-Doubling in Modern Greek. *Lingua* 114:963-989.
- Revithiadou A. & Spyropoulos V. (2008). Greek Object Clitic Pronouns: A Typological Survey of their Grammatical Properties. *Sprachtypologie und Universalienforschung* 61: 39-53.
- Ruijgh C. J. (1967). *Études sur la grammaire et le vocabulaire du grec mycénien*. Amsterdam.
- Siewierska A. (1999). From Anaphoric Pronoun to Grammatical Agreement Marker: Why Objects Don't Make It. In G. Corbett (ed.) *Agreement* (Special issue of *Folia Linguistica* 32/2), 222-251.
- Sihler A. L. (1995). *New Comparative Grammar of Greek and Latin*. New York and Oxford.
- Szemerényi O. J. L. (1996). *Introduction to Indo-European Linguistics*. Oxford.
- Taylor A. (1996). A Prosodic Account of Clitic Position in Ancient Greek. In A.M. Zwicky and A. Halpern (eds.) *Approaching Second. Second-Position Clitics and Related Phenomena*, 478-503. Stanford.
- Ventris M. & Chadwick J. (1973). *Documents in Mycenaean Greek*. 2nd ed. by J. Chadwick. Cambridge.
- Zwicky A. M. (1977). *On Clitics*. Bloomington.

Inter-dialectal Insights into Greek Rhythm: The Case of Standard Modern Greek vs. Kozani Greek

NINA TOPINTZI
*Aristotle University of
Thessaloniki & University of
the Aegean*
topintzi@enl.auth.gr

KATERINA NICOLAIDIS
*Aristotle University of
Thessaloniki*
knicol@enl.auth.gr

ELENI TSIARTSIONI
*Aristotle University of
Thessaloniki*
eltsiartsioni@hotmail.com

1. Introduction and aims

The traditional rhythm class typology categorises languages into stress-timed, syllable-timed and mora-timed (Pike 1945, Abercrombie 1967, Ladefoged 1975). This categorization is based on the notion of isochrony of the organizing units postulated for each rhythm category. For stress-timed languages, such as English and German, it is claimed that the temporal interval between stresses (or feet) is of relatively equal duration; for syllable-timed languages, such as French and Spanish, syllables are postulated to recur at relatively regular temporal intervals while for mora-timed languages, such as Japanese, there is relatively even temporal spacing between successive moras. Subsequent experimental studies have however failed to provide support for isochrony; for example, inter-stress intervals have been found to be longer when they contain more syllables, i.e. they are proportional to the number of syllables they contain. Inter-stress intervals do not appear more regularly in stress-timed compared to syllable-timed languages (Bolinger 1965, Lehiste 1977, Dauer 1983, 1987). Syllables and moras are not of nearly equal length in syllable- and mora-timed languages (Roach 1982, Dauer 1983, 1987). In view of such findings, Dauer (1983) proposed that rhythmic differences between languages are a consequence of language structure. Stress-timed languages typically have more varied syllable types and hence more variation in syllable length. They also exhibit extensive reduction phenomena in the absence of stress. Dauer (1987) presented a set of parameters that differentiate rhythm types and on the basis of these she proposed that languages are placed on a continuum from least to most stress-timed.

More recent research has introduced the use of rhythm metrics to classify languages into different rhythm types by quantifying consonantal and vocalic variability (e.g. Ramus, Nespor & Mehler 1999, Grabe and Low 2002, White and Mattys 2007). Ramus Nespor & Mehler (1999) have shown that there is greater consonantal variation in stress-timed languages while in syllable-timed languages a higher percentage of the overall utterance is vocalic. Grabe and Low (2002) using a different pair of variability indices have shown that stress-timed languages are characterized by relatively high values in vocalic and intervocalic (i.e. consonantal) intervals. This reflects variability in syllable structure, e.g. presence of clusters in onset or coda position, as well as reduced vowels in unstressed position. Syllable-timed languages on the other hand show low values for the variability indices due to the fact that they commonly have a simple CV structure and there is little vowel reduction resulting in low durational variability between successive vowels.

The results of these studies have however shown that there are several classification problems and thus limitations from the use of metrics. For instance, the metrics used by Ramus, Nespor & Mehler (1999) and by Grabe and Low (2002) classify some languages differently, e.g. Polish, Greek. In addition, while the metrics used by Grabe and Low (2002) classify appropriately languages such as English and Spanish, which have been prototypically described as stress- and syllable-timed respectively, they encounter problems with other non-prototypical languages some of which remain unclassified.

On the basis of findings from her research and a review of the results from previous studies, Arvaniti (2009) provides a detailed account of the limitations of rhythm metrics

and argues why they cannot classify languages reliably. She also discusses several factors that can influence metric scores, e.g. the choice of speech materials. She claims that metrics based on durational measurements reflect timing; this can be affected by various factors, e.g. stress, focus, context, etc. Timing relates to rhythm but it is not its exclusive component. She proposes an alternative account according to which language rhythm is based on the principles of grouping and prominence. Among other languages, she puts forth evidence from Greek to support her claims. Greek is an interesting case because it has been classified as syllable-timed, stress-timed, mixed or remained unclassified (Barry and Andreeva 2001, Baltazani 2007, Tsiartsioni 2008, Johnson & Sinabaugh 1985, Grabe & Low 2002). As mentioned above, Dauer's (1983) findings showed that the duration of inter-stress intervals was similar between stress-timed languages, such as English, and syllable-timed ones, such as Spanish. Differences were however evident in the number of syllables present in the intervals. While stresses appeared at relatively regular intervals in all languages, languages such as Spanish, Italian and Greek had more syllables than English between stresses. In line with Dauer's view that rhythm is stress-based, i.e. languages are placed on a more or less stress-timed continuum, Arvaniti claims that "one difference between languages called stress-timed and those called syllable-timed may have to do with the spacing of prominences, not in terms of duration but in terms of number syllables; in this respect, prominences may be sparser in syllable-timed languages" (2009: 59). The relative regularity in the occurrence of prominences is the result of language specific factors including reduction phenomena and speaking rate differences.

Such different approaches in the research of rhythm can have important implications for its study among different languages as well as for the cross-dialectal study of rhythm within a particular language. Variation in rhythm among dialects may actually be an important factor that contributes to their differentiation. Cross-dialectal variation in rhythm has been reported, among others, for Taiwan and American English (Jianm, (2004), Singapore and British English (Ling, Grabe & Nolan 2000), Bari, Naples and Pisa varieties of Italian (Barry, Andreeva, Russo, Dimitrova & Kostadinova 2003), European and Brazilian Portuguese (Frota & Vigario 2001), Peruvian Spanish (O' Rourke 2008), Eastern and Western varieties of Arabic (Ghazali, Hamdi & Barkat 2002), Cantonese, Beijing Mandarin, Cantonese-accented and Mandarin-accented English (Mok & Dellwo 2008).

To date, there has been no research on the rhythm of different Greek dialects. The current study investigates the speech rhythm of Kozani Greek (KG) and Standard Modern Greek (SMG). KG is a typical Northern Greek dialect displaying the raising of unstressed mid /e, o/ to high [i, u], and the deletion of unstressed underlying /i, u/ (see Dinas 2005). Vowel deletion (VD) leads to the creation of a variety of consonantal sequences which are both more numerically and longer segmentally when compared to those found in SMG. Thus, KG can present a more complex syllable structure than SMG. As noted above, complex syllable structure and vowel reduction are two factors that may affect a language's rhythmic classification. It may be expected therefore that there is variation in rhythm between KG and SMG.

Two approaches will be used in the current study: (a) rhythm metrics, in particular the vocalic and consonantal Pairwise Variability Indices (PVI) proposed by Grabe and Low (2002), and (b) inter-stress intervals (ISI), in particular the measurement of the number of syllables between stresses, in line with the prominence-based theoretical framework proposed by Arvaniti (2009). The study aims to compare and evaluate the findings from the two approaches. It also aims to look into speaker and speech material variability.

The paper is structured as follows: Section 2 presents the methodology used in the study. Section 3 presents the results of the two approaches including evidence for inter-dialectal and inter-speaker variation. Section 4 provides a discussion of the results and expounds on the basic principles of the ISI approach.

2. Methodology

2.1. Subjects

Two KG speakers (TL: male, KS: female) and two SMG speakers (TT: male, BT: female) were recorded. The speakers were between 58 to 66 years of age. The KG speakers were born, raised and lived in Kozani. Of the SMG speakers, BT was born and raised in Athens in a SMG environment. TT was born in Kozani and left for Athens at the age of 11 where he grew exclusively monodialectal in SMG. Both his parents were speakers of SMG, although his mother and other family had been exposed to KG.

2.2. Recording materials and procedure

The speech material consisted of (a) a text written in SMG which the subjects were asked to read at a comfortable speaking rate, and (b) quasi-spontaneous speech produced during a picture-description task. The text was a short narrative designed to include many potentially VD undergoing words, e.g. /skulici/ 'worm', /ɣurúɲa/ 'pigs', /ðu'la/ 'work' (see Appendix). These were expected to be realised differently by the speakers of the two dialects, i.e. [sku'lici], [ɣu'ruɲa], [ðu'la] by SMG speakers and ['sklic], ['ɣruɲa], ['ðla] by KG speakers. Picture description was chosen over entirely free speech so that data from the two dialects were more comparable, since similar vocabulary was expected to be used by all speakers. These two types of elicitation tasks were included in the present study in order to investigate possible variation in speech rhythm due to speech material (cf. Arvaniti 2009, Ferjan et al 2008, Ross et al 2008a, 2008b).

The SMG speakers were recorded in a quiet room at their home, whereas the KG speakers were recorded in a quiet room in a cultural centre. The researchers who conducted the recordings were familiar to the speakers. The subjects of both dialects were given some time before the recordings to practise reading the text and make sure they were fluent enough for the reading task. They were also given some time to look at the picture before the recordings.

2.3. Data analysis

2.3.1. PVI

The Pairwise Variability Indices (i.e. consonantal and vocalic PVIs) which express durational variability in successive vocalic and intervocalic intervals were used in this study (Grabe and Low 2002). These metrics were selected because they have been commonly used in the literature thus enabling the comparison of the results of the current study to findings from previous literature. Grabe and Low (2002: 524) define vocalic intervals as 'the stretch of signal between vowel onset and vowel offset regardless of the number of vowels included in a section' and intervocalic intervals as 'the stretch of signal between vowel offset and vowel onset, regardless of the number of consonants included'. Low PVIs indicate that variability in duration is low, i.e. the duration of successive measurements is relatively similar as expected in the co-called syllable-timed languages. In contrast, high variability indices are anticipated in stress-timed languages reflecting complex syllable structure and reduced vowels.

To compute the PVIs, consonantal and vocalic intervals were segmented using PRAAT (Boersma and Weenik 2007) in line with the aforementioned criteria as described in Grabe and Low (2002). Figure 1 illustrates segmentation of the word /skuliki/ 'worm' produced by a speaker of SMG and KG.

*Inter-dialectal Insights into Greek Rhythm:
The Case of Standard Modern Greek vs. Kozani Greek*

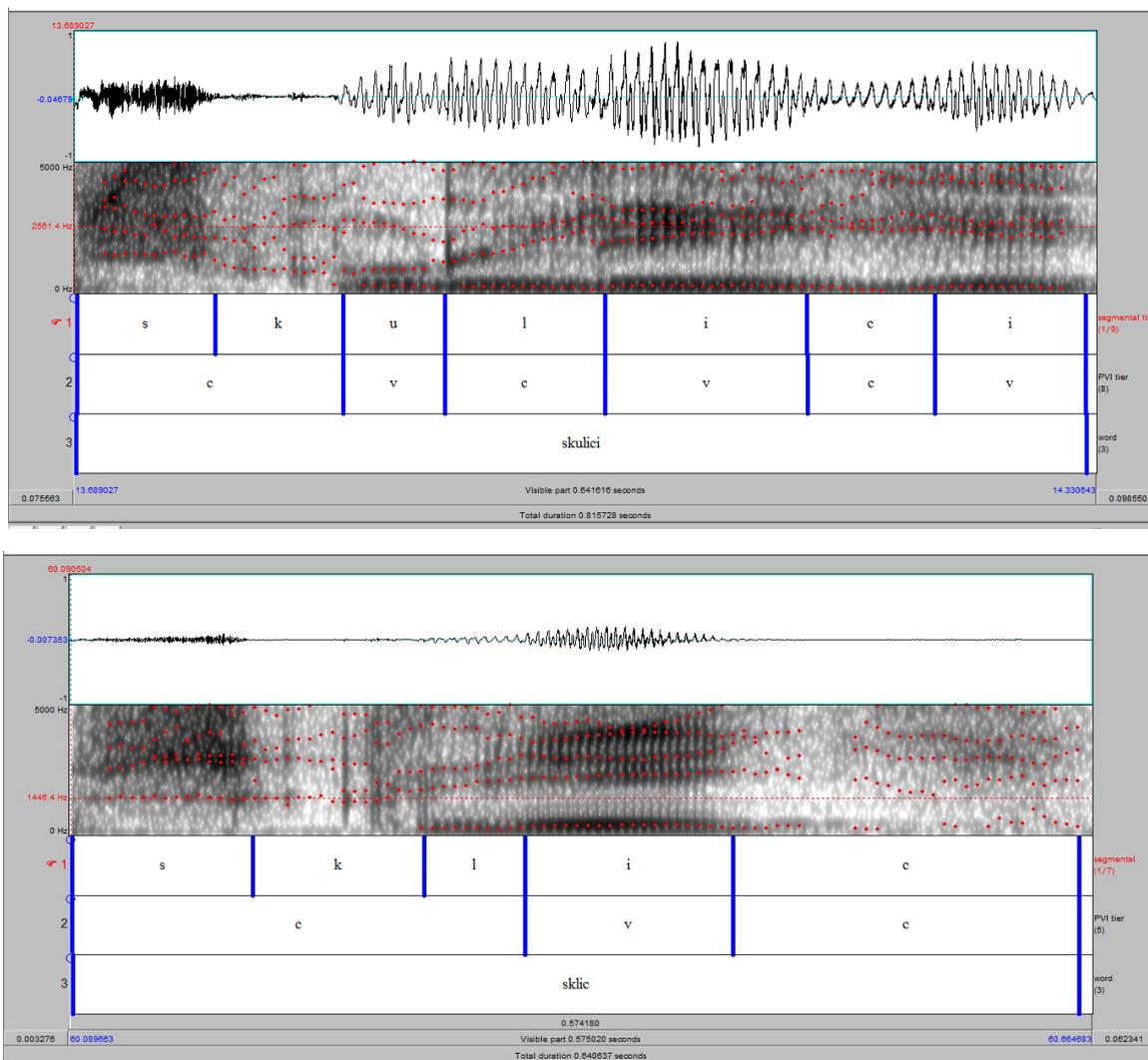


Figure 1: Segmentation of the word /skuliki/ produced as [skulici] by the SMG speaker TT (above) and as [sklic] by the KG speaker TL (below).

N.B: tier 1 = segmental tier, tier 2 = PVI tier, i.e. vocalic and consonantal intervals, tier 3 = word tier

The following segmentation principles were followed:

- (a) Pauses and hesitation marks were excluded from measurement.
- (b) Utterance-initial voiceless stops /p, t, k/ were included in the analysis. To determine their onset we estimated the average duration of all non-utterance initial and medial /p/, /t/, /k/ in the data. Thus, for each utterance initial stop, total duration was taken to correspond to the average duration of its non-utterance initial and medial counterparts.
- (c) The alveolar /r/ in consonant clusters often included a vocoid having formant structure similar to that of a short vowel (see Arvaniti 2007, Baltazani 2005, 2009 for discussion). This part was included in the consonantal interval for phonetic/phonological reasons: since rhotics are considered to be consonants in most languages, the particular vocoid was considered as part of the articulation of the rhotic sound in the specific context. Figure 2 illustrates this realisation of /r/ in the word /traɣuðai/ by the KG speaker TL.

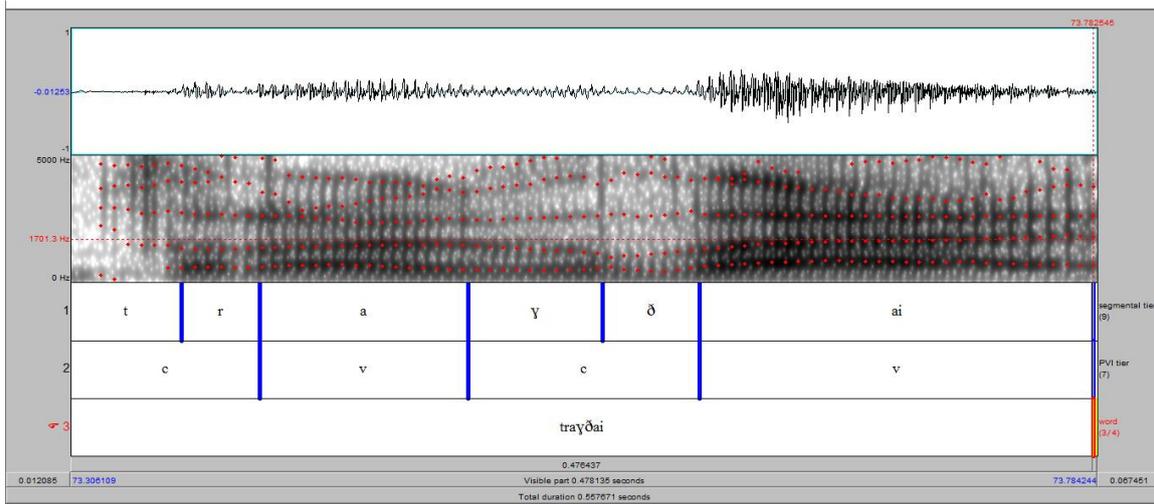
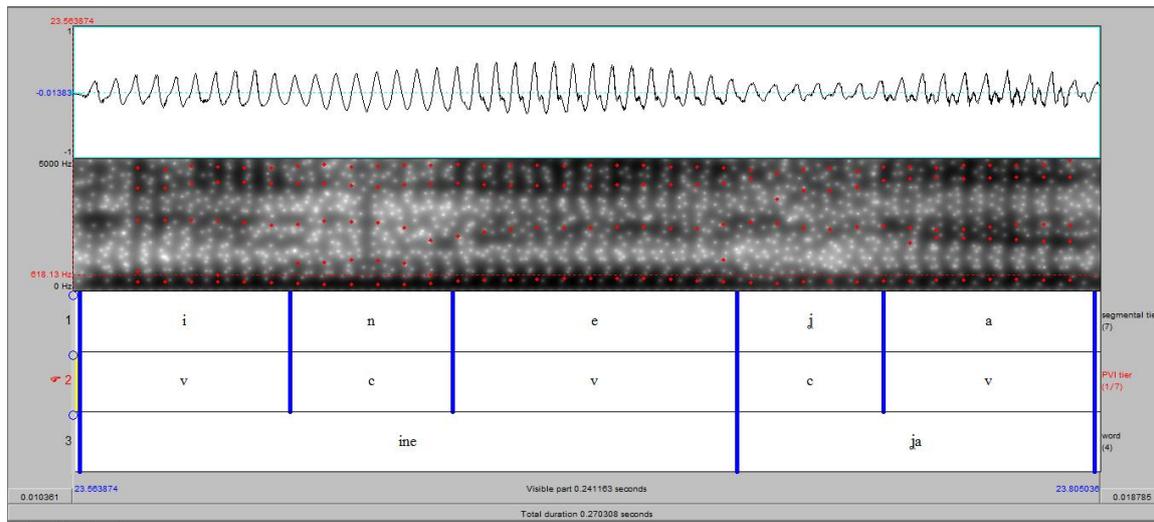


Figure 2: Spectrogram of the word /τραυδαί/ produced as [traɣðai] by the KG speaker TL.

(d) The voiced palatal fricative [j] was considered to be part of the consonantal portion when there were clearly observable changes in the formant structure or the amplitude of the signal (Figure 3 top). When reduced, i.e. articulated as an approximant, there were no clearly observable changes in the formant structure or the amplitude of the signal; in this case it was included in the vocalic portion (cf. Grabe and Low (2002) for the segmentation of glides and Malavakis 1984, Arvaniti 1999, 2007, Nicolaidis 2003, for the phonetics of Greek glides) (Figure 3 bottom).



*Inter-dialectal Insights into Greek Rhythm:
The Case of Standard Modern Greek vs. Kozani Greek*

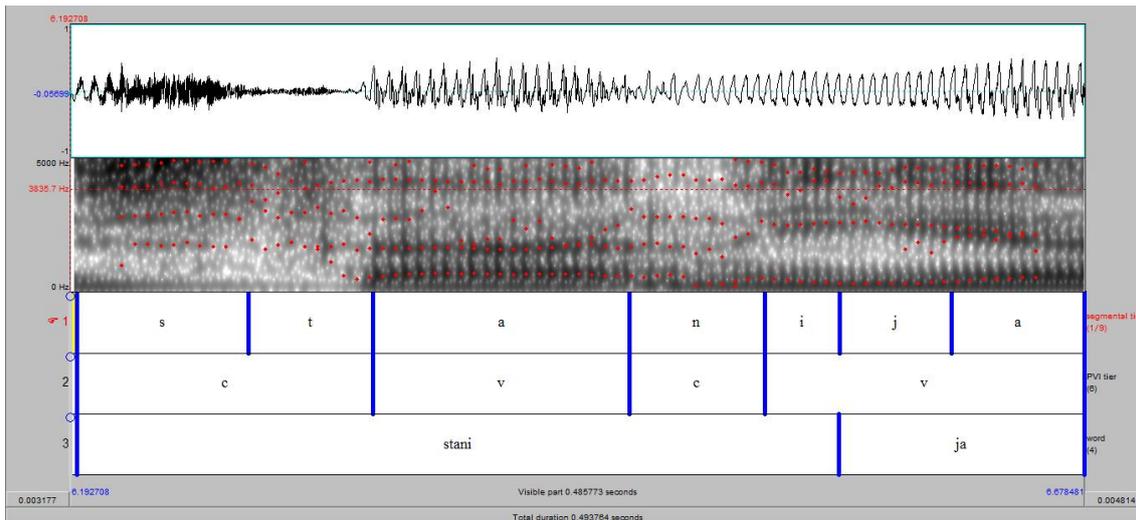


Figure 3: Spectrograms of the words [ine ja] (top) and [stani ja] (bottom) produced by the SMG speaker BT

(e) Reduced vowels were segmented as follows: those showing no evidence of voicing (i.e. no voice bar) were considered to be part of the consonantal interval (Figure 4, vowel /i/ in [astamatita]), whereas vowels with presence of voicing were considered to be part of the vocalic portion.

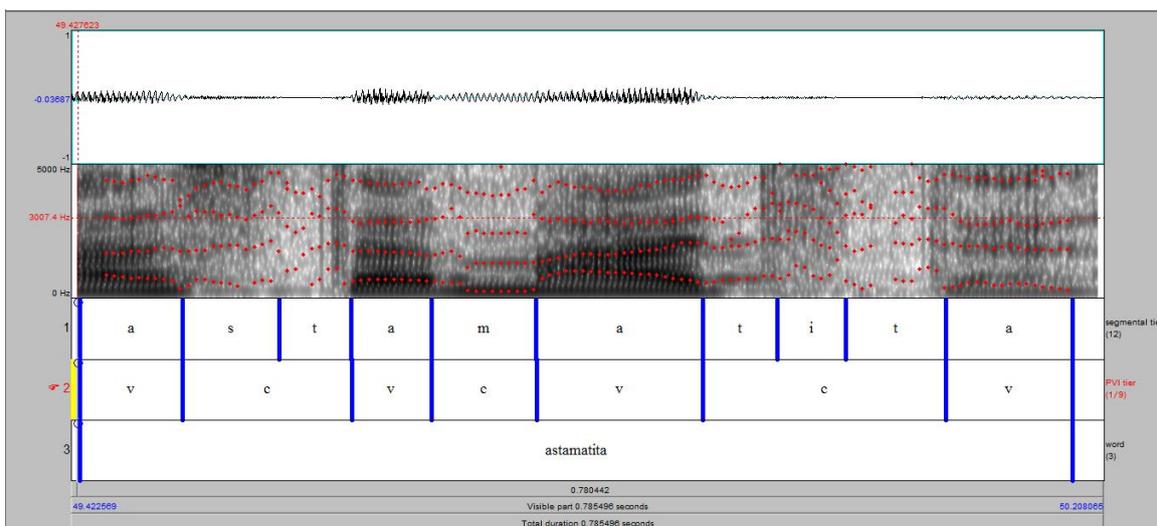


Figure 4: Spectrogram of the word /astamatita/ produced by the KG speaker KS.

An average of 129 vocalic and 126 consonantal intervals were measured for each speaker in the reading task and of 108 vocalic and 109 consonantal intervals in the quasi-spontaneous task. Subsequently a normalised version of the PVI was computed for the vocalic and consonantal intervals. The PVI is based on the mean difference in duration between successive vocalic and consonantal intervals divided by the sum of the same intervals. Normalised vocalic and intervocalic PVI were used, as normalisation adjusts for potential speaker rate variation due to the different types of the elicitation tasks (Bunta and Ingram 2007). For each passage, a PVI score for vocalic and a PVI score for consonantal intervals was computed (vocPVI and consPVI respectively). Statistical analyses were not conducted due to the small number of speakers. A detailed presentation of the results appears in section 3.1.

2.3.2. Inter-stress intervals (ISI)

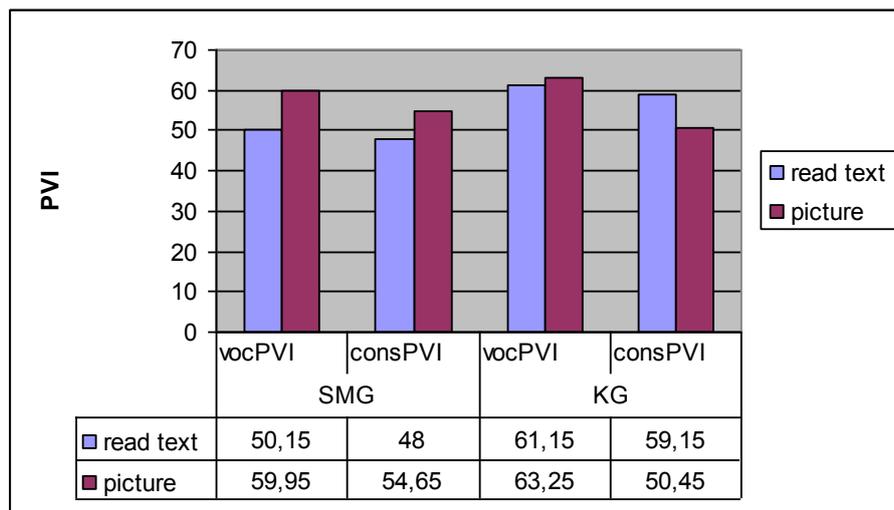


Figure 5: Average vocalic and consonantal PVI values for SMG and KG

Table 1: Vocalic and consonantal PVI scores for individual speakers in the read text and picture elicitation tasks.

STANDARD MODERN GREEK					KOZANI GREEK				
speaker	gender	material	vocPVI	consPVI	speaker	gender	material	vocPVI	consPVI
TT	male	text	48,4	46,3	TL	male	text	51,70	55,80
TT	male	picture	54,6	60,3	TL	male	picture	64,4	53,4
BT	female	text	51,9	49,7	KS	female	text	70,6	62,5
BT	female	picture	65,3	49	KS	female	picture	62,1	47,5

Figure 5 shows that KG has higher average PVI values than SMG, except for consPVI in the picture task. The high consPVI values for SMG in this task can be largely attributed to the high consonantal variability of the speaker TT (Table 1). Inter-speaker variability was also observed in the data. For example, there is a great difference in the vocPVI of the two KG speakers in the reading task and a great difference in the vocalic and consonantal PVI scores for the two SMG speakers in the picture task.

Variation due to elicitation type is also evident. For SMG, average PVI scores for the picture task are higher than those for text reading. KG shows the opposite pattern with generally higher scores for text reading than the picture task (with the exception of the voc PVI for speaker TL) (Table 1). It should be noted that the text used in this study was designed to include many words that could potentially undergo VD in KG. This is expected to result in the presence of many consonantal sequences and hence increased consonantal variability for the text-reading elicitation task. This is clearly illustrated in Figure 6 which plots consPVI against vocPVI. The KG read text appears on the right side of the graph as a result of the large consPVI. It should be noted that PVI scores closer to the upper-right side of the graph are indicative of more ‘stress-timing’, whereas scores closer to the lower-left side indicate lower variability and, thus, more ‘syllable-timing’. For SMG quasi-spontaneous speech from the picture task shows greater PVI values than the read text indicating more variability in the former task.

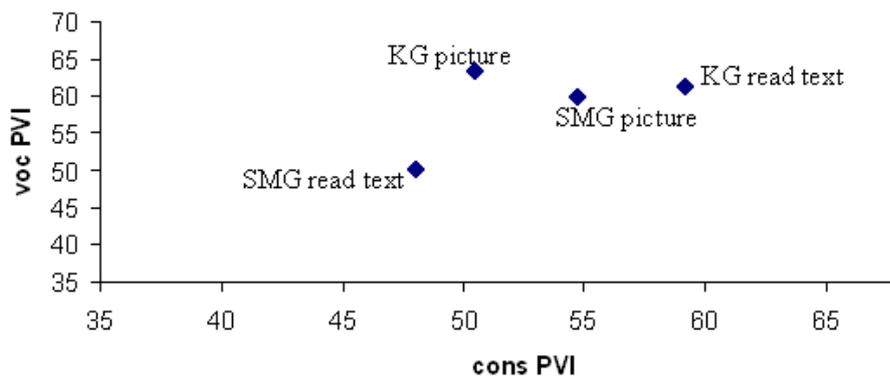


Figure 6: *The PVI profile of SMG and KG.*

3.2. Inter-stress intervals

Table 2 presents the average number of unstressed syllables between stresses for all speakers in the two elicitation tasks. On average, SMG tends to compress more unstressed syllables between stresses than KG in both text and picture, i.e. a larger number of unstressed syllables intervenes between stresses in SMG than KG. Speaker variability was however evident with KG speaker KS producing more unstressed syllables than the two SMG speakers in the picture task.

With reference to the influence of elicitation task, for SMG speakers, the averages for the read text are higher than those for the picture. This could be attributed to the presence of more instances of VD in quasi-spontaneous speech for the SMG speakers (see Dauer 1983, Baltazani 2007). Text reading may be assumed to involve a more careful style of production thus not favouring as many instances for VD as more free types of speech.

While, on these grounds, similar results may be expected for KG, i.e. more unstressed syllables in the read text than quasi-spontaneous speech, interesting speaker variation is evident in the data. Speaker TL shows a similar pattern to SMG speakers while speaker KS produces slightly fewer unstressed syllables in the read text. Overall, KS has the largest number of unstressed syllables of all speakers in the picture task. It is interesting to note that in the read text there is no large inter-speaker variation within SMG or KG indicating relatively similar speaker behaviour in the more formal style of speech. For SMG, inter-speaker variation is consistent between elicitation types, i.e. BT consistently produces fewer unstressed syllables than TT in both tasks; this can be interpreted as more instances of VD by this speaker. For KG, speaker KS consistently produces more unstressed syllables than TL in both elicitation types indicating less VD by this speaker. Evidence of more unstressed syllables in the picture task than the read text may relate to individual/idiolectal preferences during the production of freer speech, i.e. relatively careful production, or to possible planning/execution strategies, for example prolongation of unstressed vowels or other dysfluencies, which may affect the number of unstressed syllables between stresses.

Table 2: *Average number of unstressed syllables between stresses per speaker across dialects and elicitation tasks*

	SMG		KG	
	BT	TT	KS	TL
Read text	1.71	1.83	1.56	1.44
Picture	1.39	1.49	1.65	1.14

To investigate the regularity of prominences / beats, we examined the frequency of occurrence of unstressed syllables, i.e. how many 0σ , 1σ , 2σ , etc. intervals appeared

between stresses, for each speaker and text type. In figures 7 and 8, percentage data are pooled for both speakers in SMG and KG, so that dialectal differences are shown.

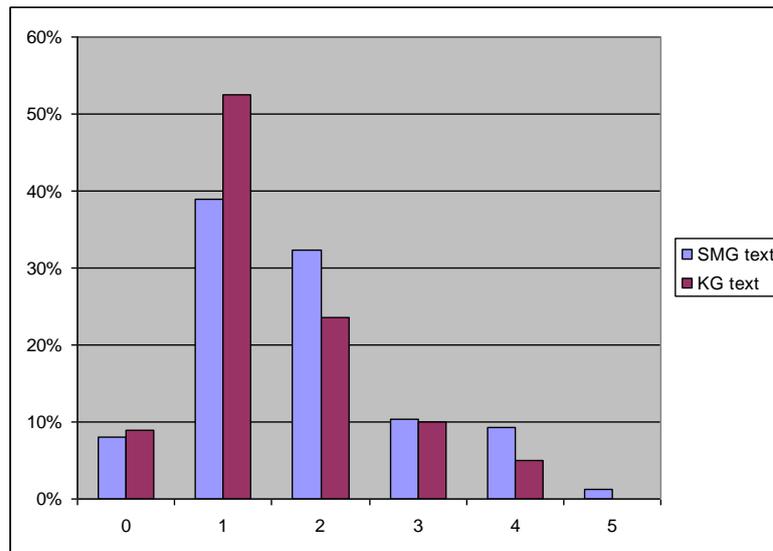


Figure 7: Percentages of unstressed syllables between stresses cross-dialectally in 'read text'

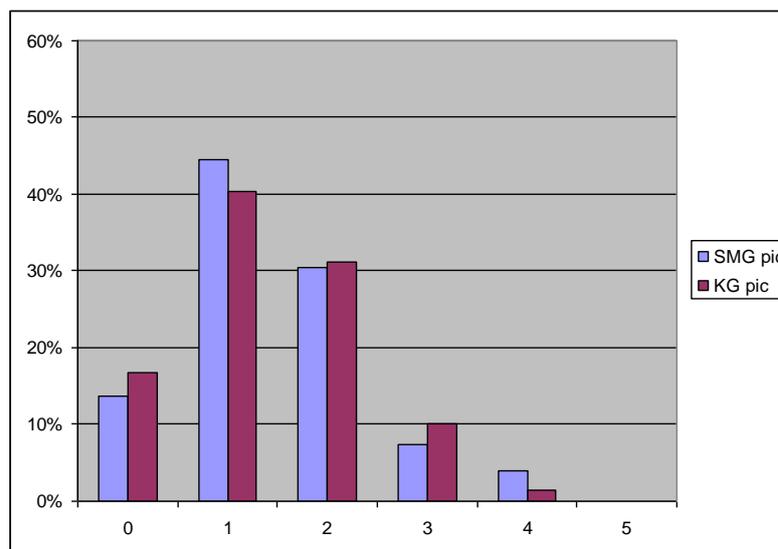


Figure 8: Percentages of unstressed syllables between stresses cross-dialectally in 'picture'

With reference to the read text, the preferred ISI is 1σ for KG; this appears over 50% of the time. The 2σ ISI shows relatively high frequency, but compared to the 1σ ISI, it appears less frequently at approximately 25%. The remaining intervals appear much less frequently, i.e. less than 10%. For SMG speakers the 1 and 2σ intervals are also preferred, but their percentages are much more balanced, i.e. 39% vs. 32% respectively. The remaining ISIs appear less frequently, less than 10%, but intervals longer than 3σ , emerge slightly more frequently in this dialect compared to KG. In fact, a single instance of a 5σ interval, the longest present in the data, is found in the read text of SMG.

With reference to the picture task, preference for 1σ and 2σ ISIs is evident in both dialects. Larger ISIs (3σ and 4σ) appear less frequently, i.e. less than 10%, in both dialects. Comparing the two elicitation types (Figures 7 and 8), an interesting increase in the 0σ ISIs is evident in the picture task for both dialects. For KG, a decrease is evident in the 1σ ISIs and an increase in the 2σ ISIs in the picture task compared to the read text.

Interesting speaker variability especially between the KG speakers was also evident in the data (Figures 9 and 10). For the read text, relatively large differences between the two KG speakers are evident for the 2σ ISIs followed by the 0σ and 3σ ISIs.

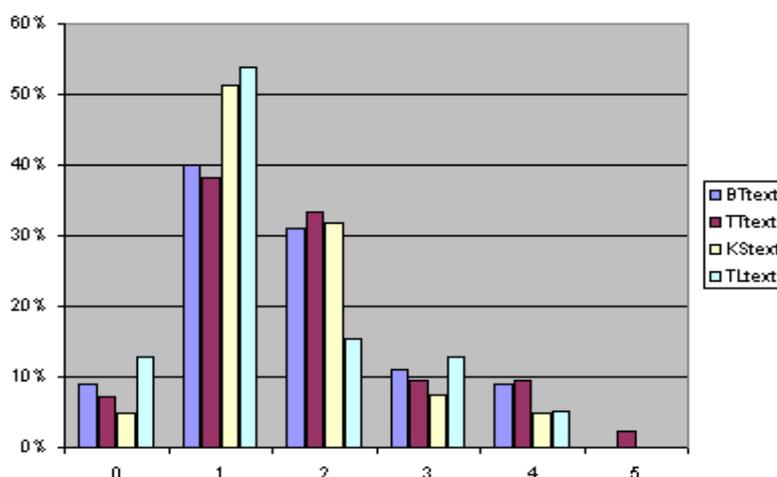


Figure 9: Percentages of unstressed syllables between stresses for each speaker in 'read text'

For the picture task, there is a very large difference between the KG speakers for the 1σ ISI which is lower by 20% for speaker KS compared to TL. Relatively large differences between the KG speakers are also evident for the 3σ ISI. In addition, KS is different from the other SMG and KG speakers in that she has a higher percentage for the 2σ than the 1σ ISI.

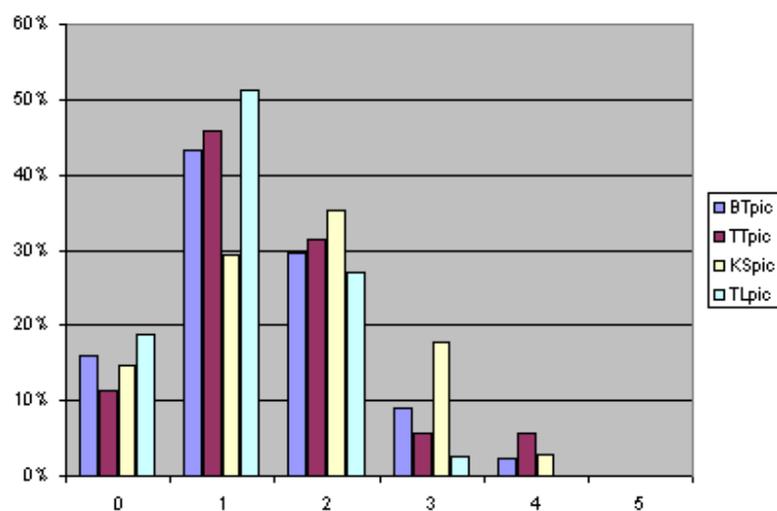


Figure 10: Percentages of unstressed syllables between stresses for each speaker in 'picture'

4. Discussion

The aim of this study has been to investigate potential differences in speech rhythm between SMG and KG. KG typically includes processes such as VD of unstressed high vowels (/i/ and /u/) and as such it is expected to present a more complex syllable structure due to the creation of consonantal sequences after the application of VD. Complex syllable structure and vowel reduction may affect its rhythmic classification differentiating it from SMG.

The results of the PVI analysis for the read text indicated that KG had more consonantal and vocalic variability and was thus more stress-timed compared to SMG. Since the text used in the current study included many potential instances of VD in KG, it

can be argued that these findings support Arvaniti's (2009) findings on the influence of type of material on metric scores. As a result of VD, the text included more instances of consonantal sequences which together with variability in the vocalic intervals resulted in the more stress-timed classification of KG.

Such influence is not expected in the quasi-spontaneous speech from the picture task. The PVI results showed larger vocPVI for KG which suggests greater vocalic variability in this dialect. This is in line with the results of Topintzi and Baltazani (to appear) who argue that VD is gradient and variable in KG. It is gradient in the sense that outputs of the process range from truly elided vowels to fully voiced ones along with various intermediate realisations, such as completely or partly devoiced. It is also variable, because tokens of the same word may sometimes undergo vowel deletion and sometimes not. While gradient and variable effects have been reported for VD in SMG by Dauer (1980) and Baltazani (2007), these results suggest greater variability in KG than SMG.

On the other hand, the consPVI was greater for SMG than KG indicating more consonantal variability in SMG. Although both SMG speakers showed this tendency, there was inter-speaker variability as one of the speakers (TT) had a very high consPVI and thus a large difference from the KG speakers. The other speaker had relatively similar consPVI with the KG speakers. Taken together, these results for the quasi-spontaneous speech (for which there is no expected imbalance towards more instances of VD for the SMG dialect) show different tendencies for KG and SMG which do not clearly differentiate the two dialects in the more or less stress-timed continuum.

The results of the ISI approach have shown that on average there are more unstressed syllables between stresses in SMG than KG. This is the case for both speakers in the read text. Such a finding may relate to the text used; many instances of VD by the KG speakers may be expected to result in fewer unstressed syllables between stresses. The results for the picture task may thus be expected to show more representative cross-dialectal differences. Speaker variation was however evident in the data. Similarly to the results for the read text, one of the speakers produced fewer unstressed syllables compared to the SMG speakers. The other speaker though produced more unstressed syllables than the SMG speakers. The results do not therefore provide conclusive evidence as to possible differences in the spacing of prominences in the two dialects. Following Arvaniti (2009), prominences may be expected to be sparser in SMG if this is assumed to be less stress-timed than KG due to simpler syllable structure and less VD effects.

In terms of the frequency of occurrence of unstressed syllables, both SMG and KG show a higher frequency of occurrence of 1σ and 2σ ISIs over less (0σ) or more (3σ , 4σ , 5σ) unstressed syllables. The highest percentages are evident for the 1σ ISIs in both dialects. Variation was evident due to elicitation type. In the read text KG speakers showed a strong tendency for the 1σ ISIs over the 2σ ISIs (twice as many 1σ ISIs), while SMG speakers had smaller differences between the 1σ and 2σ ISIs. Such a finding may relate to the read text used (see above). In the picture task, the results show similar tendencies between SMG and KG, i.e. overall highest percentage for 1σ ISIs followed by 2σ ISIs. There was however important speaker variability especially for the KG speakers. One KG speaker showed a slightly higher percentage for the 2σ ISI over the 1σ ISI. The other KG speaker showed a strong preference for the 1σ ISI over the 2σ ISI similarly to the read text. For both dialects, an interesting increase in the 0σ ISI was also evident indicating more instances of adjacent stresses (i.e. stress clashes) in the quasi-spontaneous speech.

To sum up, the above findings have provided important information regarding consonantal and vocalic variability as well as the interstress intervals in SMG and KG. Variation due to speaker and elicitation type has clearly shown that more speakers and different types of material are necessary before any conclusive evidence is provided regarding possible rhythmic differentiation between the two dialects.

Overall, we believe that the ISI approach adopted in this study may provide a promising line of research in the study of rhythm in line with Arvaniti's (2009) suggestion

that rhythm should be dissociated from timing following a distinction used in psychology, namely that "timing is concerned with the durational characteristics of events, while rhythm has to do with the pattern of periodicities that is extracted from these durations" (2009: 59). This approach also offers a tentative interpretation – by means of interstress intervals – of Arvaniti's statement that rhythm is created on the basis of grouping and prominence patterns. In particular, we have set aside reference to consonantal and vocalic variability and speculate that rhythm primarily relies on two dimensions:

- (a) Beat Frequency
- (b) ISI distribution

Beat frequency refers to the preferable spacing between beats expressed as the mean value of the ISIs. The higher this number, the more syllables are clustered between stresses or, alternatively, the longer the ISI in terms of syllables. ISI distribution refers to the distributional pattern of unstressed syllables, i.e. whether most ISIs are clustered together in a certain area, e.g. that of $1-2\sigma$, or they are more evenly distributed among various syllables, e.g. roughly equally spread between $0-4\sigma$.

Several aspects of this approach need to be fine-tuned and evaluated in future work, e.g. methodological issues such as the exclusion of unstressed syllables following a pause and theoretical issues such as the influence of vowel/consonant length and mora structure on overall rhythm. While the results of the current study have not been conclusive for the two dialects of Greek (SMG and KG) – possibly because they are not very different – it will be interesting to test the ISI account in other languages, especially those that are traditionally termed stress-timed (English, Dutch) and syllable-timed (Spanish, French). Another testing ground is languages with fixed primary stress and no secondary stress even in polysyllabic words of 5 or 6 syllables, e.g. Mohawk (Michelson 1989, see also Heinz 2007), especially if they lack processes of vowel reduction/deletion while speaking rate is not very high. The ISI approach predicts that ISI distribution should be quite scattered, because in the absence of secondary stress, the spacing between fixed stresses will be relatively variable, thus creating the percept of less 'canonical' rhythm.

Appendix

A) Text

Κάθε μέρα με το που χαράζει ο Νίκος, νυσταγμένος, βγαίνει από το σπίτι. Περπατά αργά ως τη στάνη για να ταΐσει τα γουρούνια. Πάντα χαζεύει γύρω του. Χαζεύει το πουλί που τσιμπολογά το σκουλήκι, το χιόνι που πέφτει στα βουνά ασταμάτητα, το σκυλί που το κόκκαλο χώνει βαθιά στη γη. Το σκυλί μόλις τον βλέπει, κουνά την ουρά, τρέχει και τον φιλά. Είναι λες και του τραγουδά. Μέχρι να γυρίσει πίσω είναι για τα καλά ξύπνιος. Κάθε μέρα η ίδια δουλειά.

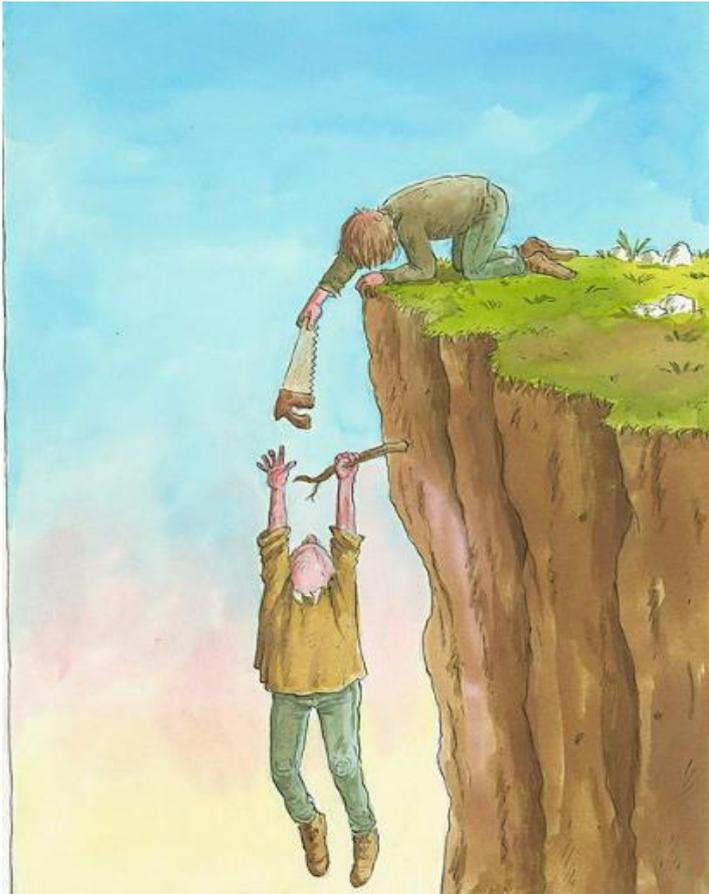
IPA Transcription (assuming SMG careful speech)

|| 'kaθe 'meɾa me to pu xa 'ɾazi | o 'nikos | nistaɣ 'menos | 'vɟeni apo to 'spiti || peɾpa 'ta aɾ 'ɣa os ti 'stani ɟa na ta 'isi ta ɣu 'ɾuɲa || 'paⁿda xa 'zevi 'ɟiɾo tu || xa 'zevi to pu 'li pu tsi^mbolo 'ɣa to sku 'lici | to 'ɟoni pu 'pefti sta vu 'na asta 'matita | to sci 'li pu to 'kokalo 'xoni va 'θɟa sti ɟi || to sci 'li 'molis ton 'vlepi | ku 'na tin u 'ɾa | treɟi | ce ton fi 'la || 'ine les ce tu trayu 'ða || 'mexɾi na ɟi 'risi 'piso | 'ine ɟa ta ka 'la 'ksipnos || 'kaθe 'meɾa i 'iðɟa ðu 'la ||

Translation

Every day as soon as it dawns, sleepy Nikos gets out of the house. He walks slowly to the pen to feed the pigs. He always looks around. He stares at the bird that pecks a worm, the snow that endlessly falls on the mountains, the dog that buries a bone deep in the ground. The dog, as soon as it sees him, wags its tail, runs towards him and kisses him. It's as if it sings to him. Until he (Nikos) comes back is wide awake. Every day the same thing goes on.

B) Picture (from: <http://www.humor-kamensky.sk/indexuk.htm>)



References

- Asu E. L. & Nolan F. (2006). Estonian and English rhythm: a two-dimensional quantification based on syllables and feet. In Proceedings of Speech Prosody 2006, Dresden, Germany
- Abercombie D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Arvaniti A. (1999). Illustrations of the IPA: Modern Greek. *Journal of the International Phonetic Association* 19: 167-172.
- Arvaniti A. (2007). Greek phonetics: State of the art. *Journal of Greek Linguistics* 8: 97-208.
- Arvaniti A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica* 66: 46-63.
- Baltazani M. (2005). Phonetic variability of the Greek rhotic sound. Poster presented at the *Phonetics and Phonology in Iberia (PaPi 05)*, Barcelona, Spain.
- Baltazani M. (2007). Prosodic rhythm and the status of vowel reduction in Greek. *Selected Papers on Theoretical and Applied Linguistics from the 17th International Symposium of Theoretical and Applied Linguistics*, Thessaloniki, 31-43.
- Baltazani M. (2009). Acoustic characterization of the Greek [r] in clusters. *Proceedings of the 18th International Symposium on Theoretical and Applied Linguistics, Salonica 4-6 May 2007*, Volume 1: 87-95
- Barry W. & Andreeva B. (2001). Cross-Language similarities and differences in spontaneous speech patterns. *Journal of International Phonetics Association* 31: 51-66.
- Barry W. J., Andreeva B., Russo M., Dimitrova S. & Kostadinova T. (2003). Do rhythm measures tell us anything about language type? *Proceedings 15th ICPHS*, Barcelona, pp. 471-474.
- Boersma P. & Weenink D. (2007). Praat. Doing phonetics by computer (version 4.6.01), <http://www.praat.org/>
- Bolinger D.L. (1965). *Form of English: Accent, Morpheme, Order*. Cambridge Mass: Harvard University Press.
- Bunta F. & Ingram D. (2007). The acquisition of speech rhythm by bilinguals Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language and Hearing Research* 50: 999-1014
- Dauer R. M. (1980). The reduction of unstressed high vowels in Modern Greek. *Journal of the International Phonetics Association* 10: 11-27

- Dauer R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.
- Dauer R. M. (1987). Phonetic and Phonological components of language rhythm. *Proceedings of the 11th International Congress of Phonetic Sciences* 5: 447-450.
- Dinas K. (2005) *The dialect of Kozani [Το γλωσσικό ιδίωμα της Κοζάνης]*. Kozani: Institute of Book and Reading.
- Ferjan N., Ross T. & Arvaniti A. (2008). L2 speech and rhythm metrics. Paper presented at ASA '08, Paris, June 29- July 4 2008.
- Frota S. & Vigario M. (2001). On the correlates of linguistic rhythmic distinctions: The European/Brazilian Portuguese case. *Probus* 13: 247-275.
- Ghazali S., Hamdi R. & Barkat M. (2002). Speech rhythm variation in Arabic dialects. *Speech Prosody*, Aix-en-Provence, France, 331-334.
- Grabe E. & Low E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology 7*, Berlin: Mouton de Gruyter, 515-546.
- Heinz J. N. (2007). *Inductive learning of phonotactic patterns*. PhD Dissertation, UCLA.
- Jianm H.-L. (2004). On the syllable timing in Taiwan English. *Speech Prosody*, Narra, Japan, 247-250.
- Johnson K. & Sinabaugh B. (1985). The simplification of the Greek vowel system. *CLS* 21/1: 189-198.
- Ladefoged P. (1975). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.
- Lehiste I. (1977). Isochrony reconsidered. *Journal of Phonetics* 5: 253-263.
- Ling L. E., Grabe E. & Nolan F. (2000). Quantitative characterizations of speech rhythm: Syllable timing in Singapore English. *Language and Speech* 44/3: 377-401.
- Malavakis T. I. (1984). Φωνηεντικές Συνέχειες: Διφθογοποίηση, Ουρανικοποίηση και Φωνηματική Κατάταξη τους [Vowel Sequences: Diphthongization, Palatalization and Phonemic Classification]. *Studies in Greek Linguistics* 4: 1-16
- Michelson K. (1989). Invisibility: Vowels without a timing slot in Mohawk. In D.B. Gerdts & K. Michelson (eds.), *Theoretical Perspectives on Native American Languages*. New York: SUNY Press, 38-69.
- Mok P. P. K. & Dellwo V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. *Speech Prosody 2008*, Campinas Brazil, 423-426.
- Nicolaidis K. (2003). Μια Ηλεκτροπαλατογραφική Μελέτη των Ουρανικών Συμφώνων της Ελληνικής [An Electropalatographic Study of the Palatal Consonants of Greek]. *Σύγχρονες Τάσεις στην Ελληνική Γλωσσολογία* [Contemporary Trends in Greek Linguistics], D. Theophanopoulou-Kontou, C. Lascaratou, M. Sifianou, M. Georgiafentis & V. Spyropoulos (eds.), Athens: Patakis, 108-127
- O' Rourke E. (2008). Speech rhythm variation in dialects of Spanish: Applying the pairwise variability index and variation coefficients to Peruvian Spanish. *Speech Prosody 2008*, Campinas Brazil, 431-434.
- Pike K. (1945). *The Intonation of American English*. University of Michigan: Ann Arbor Press.
- Ramus F., Nespors M. & Mehler J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-292.
- Ross T., Ferjan N. & Arvaniti A. (2008a). Speech rhythm and its quantification in L1. Poster presented at the 1st Southern California Workshop on Phonetics / Phonology (SCOPHO), November 1, 2008.
- Ross T., Ferjan N. & Arvaniti A. (2008b). On quantifying speech rhythm. Poster presented at the Workshop on Empirical Approaches to Rhythm, 28 March 2008, UCL, London.
- Topintzi N. & Baltazani A. (to appear). The acoustics of high-vowel loss in a Northern Greek dialect and typological implications. P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, and B. Kühnert (eds) *Clusters and Structural Complexity*. Interface Explorations Series, Berlin: Mouton de Gruyter.
- TsiartSIONI E. (2008). A PVI analysis of Greek speech rhythm among children, teenagers and adults. *Proceedings of the 29th Annual Meeting of the Department of Linguistics*, Aristotle University of Thessaloniki 10-11/5/2008, 541-552.
- White L. & Mattys S. (2007). Calibrating rhythm: first and second language studies. *Journal of Phonetics* 35 / 4: 501-522.

“The importance of being voiced”: cluster formation in dialectal variants of Greek

MARINA TZAKOSTA
University of Crete
 martzak@gmail.com

1. Introduction

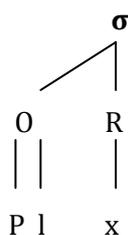
Cluster formation is one aspect of examining a language’s prosodic phonology. Variety in cluster formation is widely attested in developmental data of language acquisition (L1) and language learning (L2) as well as a language’s standard form and its dialectal variants (cf. Tzakosta 2004 and more references therein). The basic principles underlying cluster formation is that, first, the sonority scale (hereafter SS) needs to be satisfied in a rightward direction, i.e. cluster members have to be selected from left to right so that segments rise in sonority, and, second, the bigger the distance between the members of a cluster on the SS is the better structured this cluster is (Clements 1984, 1988). To give an example, /kl/ is a better formed cluster compared to /kn/ because of the bigger distance holding between /k/ and /l/ (4) as opposed to /k/ and /n/ (3). This is the reason why CL⁷² rather than CC clusters are considered to be perfect and, consequently, they emerge more frequently in various aspects of a language and cross-linguistically. The classical sonority scale is depicted in figure 1.

S	F/Sib	Affr	N	L	G	V
1	2	3	4	5	6	7

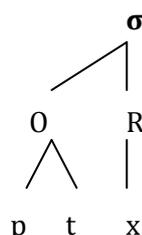
Figure 1: *The sonority scale (Selkirk 1982, Steriade 1984)*

Given the above assumptions, it is easy to gather that from a phonological, a phonetic and a psycholinguistic point of view, a well-formed, and, ideally, a perfect cluster has more chances to remain intact in its surface/ phonetic realization. In other words, big sonority distance among cluster members leads to ‘clear’ cluster perception; in turn, this clarity drives easy production. Tzakosta & Vis (2009) reach the same conclusions based on developmental L1 data. They argue that the phonetic gap existing between members of a perfect cluster facilitates perception and production. The smaller the distance between cluster members the more difficult it is for these clusters to be accurately perceived and produced. The theoretical connotation of this claim is that different clusters are characterized by different phonological representations. Perfect CL clusters tend to be characterized by a more ‘loose’ phonological representation, whereas CC clusters are characterized by a more coherent and ‘tight’ representation. These theoretical differences are demonstrated in schemas (a) and (b) in figure 2 below, respectively.

a. CL clusters



b. CC clusters



⁷² C stands for obstruent consonants (stops and fricatives), L for liquids and rotics, N for nasals.

(cf. Tzakosta 2009, Tzakosta & is 2009)

Figure 2: *Differences in the phonological representations of different consonant clusters*

However, dialectal data illustrate that non-perfect clusters not existing in the standard language emerge in dialectal variants. Such non-perfect clusters may be acceptable or non-acceptable/ non-existing. In the remainder of the paper, by non-acceptable/ non-existing clusters we will be referring to consonantal sequences disallowed by the phonotactic constraints of the standard language. In Greek, for example, sequences /pk/, /tk/ and /pθ/ do not emerge in the norm - being mainly, the result of vowel loss -, though they emerge massively in dialects of both the northern and southern dialectal zone. We assume that this flexibility is attributed to the fact that dialects are less strict regarding their phonotactic constraints. The major characteristic of acceptable and non-acceptable/ non-existing clusters is that their members are highly adjacent on the SS, and, therefore, they do not rise in sonority. As a result, non-existing clusters are highly coherent; consequently, they are not easy to be perceived and produced.

Recent studies have proposed more refined factors that determine cluster well-formedness except for the rightward satisfaction of the SS and sonority distance. More specifically, Tzakosta & Karra (in press) suggested, based on indexed dialectal data from all major dialectal zones of northern and southern Greece, that the SS needs to be distinguished in two scales, a scale representing place of articulation (PoA) and a scale corresponding to manner of articulation (MoA). These two scales facilitate a thorough and an in depth assessment of cluster well-formedness. Tzakosta & Karra (in press) demonstrated that this scale distinction succeeds in providing a more detailed and accurate description of the phonotactics of (a) perfect clusters, like /kl/ and /pl/, (b) acceptable clusters, like /vy/ and /fθ/, and (c) non-acceptable clusters, like /tf/ and /tk/. In sequence to the above, Tzakosta & Karra claim that the (vacuous) satisfaction of the PoA and MoA scales leads to distinct degrees of well-formedness. More specifically, if clusters *satisfy* the scales of *both* manner and place, they are perfect. If they *respect* the sonority of *either* place or manner, they form *acceptable clusters*, whereas if clusters *do not respect at least one* scale, they constitute 'wrong' i.e. *non-acceptable clusters*. In other words, perfection or (non-)acceptability in cluster formation is an example of gradient satisfaction of the MoA and PoA scales. The prediction following the above claims is that clusters non-existing in the standard language may emerge in dialectal variants as long as they are theoretically acceptable. However, wrong clusters are not expected to emerge.

In this paper, we add to the above claims by challenging the role of voicing in cluster formation. We prove that the dissatisfaction of the suggested voicing scale is enough for a cluster to be characterized as non-acceptable. The rest of the paper is organized as follows; section 2 elaborates on the idea of Tzakosta & Karra (in press) discussing some representative results. Section 3 develops this idea by proposing that voicing should also comprise a distinct scale which evaluates cluster well-formedness on a par with the place and manner scales. Finally, section 4 concludes the paper and poses issues for future research.

2. The linguistic evidence

Let us now turn to the data that provide evidence for the claims promoted above. The data in (1) give some representative examples of the clusters which are attested both in standard Greek and in indexed dialectal data from the major Greek dialectal zones, namely the dialects of Northern Greece (e.g. Epirus, Meleniko, Lesvos, Pontos, Thassos, Corfu, Thessalia, Kozani, Trikala, Samothraki, Thessaloniki) and of Southern Greece (e.g. Cyprus, Crete, Dodekanese, Ikaria). The examined clusters are the major CL and CC types. The data in (1) present the possible Greek cluster combinations. More specifically, except for well-formed CL and CR sequences, [voiceless stop + voiceless stop], [voiceless stop + voiceless/

voiced fricative], [voiced fricative + voiced fricative], [voiceless fricative + voiceless fricative], [voiceless fricative + voiceless stop] clusters are allowed, as shown in (1c). Interestingly, [voiced stop + voiced stop], [voiced obstruent + voiceless obstruent] and [voiceless obstruent + voiced obstruent] clusters are not attested in Greek, except for CN clusters.

- (1a) **CL** => **aplós, yláros**,⁷³
- (1b) **CR** => **ákri, éθrios**
- (1c) **CC** => **aktí, optikós, téfxos, xθés, fθinós, vyázo, avyó, ékθesi, ékδosi, péfko, xtízo**
- (1d) **CJ** => **δjo, áδjos**
- (1e) **CN** => **akmí, éθnos**
- (1f) **NN** => **amnesia**

In (2) we provide some more dialectal data. The clusters in (2a-b) are attested only in dialectal data, whereas (2c) emerges in dialectal data but only in one word of the standard language, namely ‘ατθίς’. Finally, /tθ/, in (2d), being a perfect cluster which satisfies sonority rising and sonority distance, occurs both in the norm and the dialects because it is a perfect cluster.

- (2a) /ku.fá.θi.ce/ → [k**f**á.θce] ‘become deaf – 3SG. PAST’ (Thessalia, Tzartanos 1909)
- (2b) /pi.θa.mí/ → [p**θ**a.mí] ‘span-FEM.NOM.SG.’ (Thessalia, Tzartanos 1909)
- (2c) /tu.fé.ci/ → [t**f**é.ci] ‘gun-NEUT.NOM.SG.’ (Thessalia, Tzartanos 1909)
- (2d) /ma.θé.no/ → [ma.t**θ**é.no] ‘learn-1SG.PRES.’ (Cyprus, Kodosopoulos 1994)
- (2e) /tu.lú.pa/ → [t**l**ú.pa] ‘wool-FEM.NOM.SG.’ (Margariti-Roga 1990-1991)

According to the theoretical proposal made by Tzakosta & Karra (in press), the type of clusters which emerge massively in dialectal data but not in the norm are the ones they call *acceptable clusters*. As already mentioned, this preference for acceptable rather than perfect clusters is attributed to the fact that dialects are more flexible regarding their phonotactics compared to the standard language. Moreover, acceptable clusters do not demand absolute scale satisfaction. However, the massive surface realization of acceptable clusters led Tzakosta & Karra to the assumption that the SS does not suffice in evaluating cluster well-formedness. They suggested a refined version of the sonority scale signaled by two distinct scales, the scale of PoA and the scale of MoA. Depending on the degree of satisfaction of these two scales, clusters are perfect, acceptable or non-acceptable. Figures 3 and 4 depict the scales of PoA and MoA, respectively. Like in the case of the classical SS in figure 1, both scales are satisfied as long as the selection of cluster members is rightward. Cluster well-formedness also depends on distance; the bigger the distance between cluster members the better-formed the cluster. To give some examples, /kt/ is a better cluster compared to /pt/ on the place scale. The distance of the /kt/ cluster members is 2, while the distance for /pt/ is 1. On the other hand, /kl/ is considered a better cluster compared to /xl/ on the manner scale because the distance for /kl/ is 4 whereas the distance for /xl/ is 3.

Velars	Labials	Coronals
1	2	3

Figure 3. The PoA scale

⁷³ In this set of examples, **C** stands for an obstruent, i.e. a voiceless or voiced stop or fricative, **L** represents a liquid, **R** stands for a rhotic, **J** represents a glide and **N** is a nasal.

S	F(/Sib)	Affr	N	L	G	V
1	2	3	4	5	6	7

Figure 4: The MoA scale

Such data further display that cluster well-formedness - no matter whether it refers to perfection or acceptability - is a gradient phenomenon. In other words, /kl/ is a 'better formed cluster' than /xl/ because of the bigger distance among the members of /kl/. Another important contribution of Tzakosta & Karra's theoretical proposal is that scales are vacuously satisfied in case cluster members share the same PoA, as in /pf/ or /tθ/ and/ or MoA, as in /pt/ or /xθ/.

Cluster perfection and (non-) acceptability are further schematized in tables 1-3. More specifically, tables 1 and 2 illustrate the sets of perfect, acceptable and non-acceptable clusters at the level of PoA and MoA, respectively. In both tables, we observe that clusters whose members are selected from left to right, with a relative distance among them, are perfect. To give an example, LAB + COR, VEL + LAB and VEL + COR are perfect clusters at the level of place of articulation, whereas STOP + L, FRIC + L, STOP + FRIC are perfect clusters at the level of manner of articulation. It is important to mention that the scales are vacuously satisfied when cluster members share the same PoA and/or MoA. As a result, LAB + LAB, VEL + VEL, COR + COR clusters are acceptable at the level of PoA and STOP + STOP, FRIC + FRIC and AFFR + AFFR clusters are acceptable at the level of MoA. Leftward selection of cluster members leads to the formation of non-acceptable clusters. Therefore, STOP + VEL and LAB + VEL are non-acceptable with respect to PoA and FRIC + STOP or FRIC + AFFRIC are non-acceptable clusters regarding MoA.

Clusters are acceptable under three conditions: a) if they satisfy one of the two scales and vacuously satisfy the other, b) if they vacuously satisfy both scales, and c) if they satisfy one but violate the other scale. In table 3, which displays the combined effects of tables 1 and 2, all acceptable clusters are written in square brackets. Clusters appearing in white backgrounds emerge both in standard Greek as well as its dialects; whereas acceptable clusters appearing in grey backgrounds emerge only in dialectal data. Underlined acceptable clusters appearing in white backgrounds signal rarely emerging clusters.

Vacuous satisfaction of one of the scales of PoA or MoA and violation of the other is a sufficient criterion in order to characterize a cluster as non-acceptable. Non-acceptable clusters may also violate both scales. The latter are the worst among non-acceptable clusters. This fact further supports the notion of gradience in cluster well-formedness. Gradience appears at all levels of cluster well-formedness, i.e. perfect, acceptable and non-acceptable clusters. Non-acceptable clusters appear in brackets in table 3. Underlined non-acceptable clusters emerge in morpheme boundaries, whereas non-acceptable clusters appearing in grey backgrounds emerge only in dialects.

The difference between acceptable and non-acceptable clusters is in most cases very subtle. This observation supports the claim that, not only are clusters gradient regarding the category they belong to, i.e. whether they are perfect, accept and non-acceptable; gradience characterizes each level of acceptability. In other words, there are perfect clusters which are 'better' than other perfect clusters or clusters which are more acceptable than other acceptable clusters. In addition to that, there are clusters which are, as already mentioned, the worst among non-acceptable clusters.

Table 1: Gradience in cluster formation (PoA)

Types	Perfect	Accept	Non-accept
Lab + Lab		√ /pf/	
Lab + Cor	√ /pt/		

*“The importance of being voiced”:
cluster formation in dialectal variants of Greek*

Lab + Vel			√ /pk/
Cor + Cor		√	
Cor + Lab			√
Cor + Vel			√
Vel + Vel		√	
Vel + Cor	√		
Vel + Lab	√		

Table 2: *Gradience in cluster formation (MoA)*

Types	Perfect	Accept	Non-accept
Stop + L	√ /pl/		
Fric+ L	√		
Stop + Stop		√ /pt/	
Fric + Fric		√	
Stop + Fric	√		
Fric + Stop			√ /ft/
Stop + Affr	√		
Affr + Stop			√
Fric + Affr	√		
Affr + Fric			√

Table 3 combines the effects of tables 1 and 2 and displays the sets of perfect, acceptable and non-acceptable clusters. Perfect clusters clearly satisfy both scales of PoA and MoA; they emerge in parentheses in table 3. Perfect clusters appearing in white backgrounds are clusters which emerge both in standard Greek as well as its dialects, whereas perfect clusters appearing in grey backgrounds emerge only in dialects. Perfect clusters appearing in angle brackets emerge only in morpheme boundaries.

Table 3: *Gradience in cluster formation (combined)*

Types	Stop Lab	Fric Lab	Stop Cor	Fric Cor	Stop Vel	Fric Vel	L
Stop Lab	GEM	√ (pf)?	√ (pt)	√ (pθ)!!!	√ (pk)	√ (px)	√ (pl)
Fric Lab	√ (fp)	GEM	√ (ft)	√ (fθ)	√ (fk) ?	√ (fx)	√ (fr)
Stop Cor	√ (tp)	√ (tf)	GEM	√ (tθ)?	√ (tk)	√ (tx)	√ (tr)
Fric Cor	√ (θp) ?	√ (θf)	√ (θt)	GEM	√ (θk) ?	√ (θx)	√ (θl)

Stop Vel	√(kp)	√ (kf)?	√(kt)	√ (kθ)?	GEM	√ (kx)	√ (kr)
Fric Vel	√(xp)	√(xf)	√(xt)	√ (xθ)	√ (xk)	GEM	√ (xl)

3. The current proposal

The theoretical claims discussed in this section develop the ideas promoted in the previous section and are supported by the same indexed data set from dialects of northern Greece (Epirus, Meleniko, Lesvos, Pontos, Thassos, Corfu, Attica, Thessalia, Kozani, Trikala, Samothraki, Thessaloniki, Koutsovlahika) and southern Greece (Cyprus, Crete, Dodekanese, Ikaria) investigated by Tzakosta & Karra (in press). In the present study, following the line of Tzakosta & Karra (in press), we focus on CL and CC clusters but we do not consider CJ clusters because we believe that [j] is the product of vowel raising. More specifically, the goals of the paper are, first, to discuss the surface realization of CL and CC clusters in Greek dialects, second, to investigate whether clusters have the same ‘survival chances’ across dialects, third, to evaluate the ‘importance’ of voicing in cluster formation and, finally, to make a typological account of the ‘strength’ of CL and CC clusters with respect to the three dimensions of place, manner and voicing.

Voicing has primarily been dealt with at the level of voicing vs. devoicing alternations (cf. Oostendorp 2004, 2006, among others) and to the extent (de)voicing is involved in assimilatory processes (cf. Al-Ahmadi Al-Habi to appear, Arvaniti 1999, Baroni 1997, Grijzenhout 2000). At the theoretical level, voicing has been accounted for in OT terms by means of the *NC, ND, *ND constraints (cf. Borowsky 2000, Grijzenhout 2000, Lombardi 1996, 1999, Pater 1999). The current research questions are related to the following: a) if voice assimilation applies to non-adjacent consonants, b) if voice assimilation applies within consonant clusters, and, c) if [-voi] + [+voi] clusters, like /kδ/ are acceptable, given that [+voi] + [-voi] clusters, like /δk/, are not acceptable at least in Greek.⁷⁴

The hypothesis underlying the current theoretical proposal is that, in addition to the PoA and MoA scales, the dimension of voicing should also be considered in cluster formation. In other words, all three dimensions of PoA, MoA and voicing need to be taken into account. More specifically, the PoA scale which corresponds to the fixed place hierarchy (cf. Prince & Smolensky 1993) and the MoA scale, which roughly corresponds to the classical sonority, as already proposed by Tzakosta & Karra (in press). In this paper, we propose the introduction of the voicing scale which completes cluster well-formedness. Before we elaborate on this idea, let us turn to some representative examples. The data in (3) and (4) display cases in which clusters emerge either due to consonant medial vowel loss, as shown in (3b-f) and (4a-e), or due to intra-dialectal, social or stylistic reasons, as demonstrated in (3a).⁷⁵

- (3a) /cí.pos/ → [cí.pfos] ‘garden-MASC.NOM.SG.’ (Cyprus, Kodosopoulos 1994)
 (3b) /ta.ra.tu.ró.pi.ta/ → [ta.ra.tu.ró.pta] ‘pie- FEM.NOM.SG.’ (Thassos, Tombaidis 1967)
 (3c) /pi.θa.mí/ → [pθa.mí] ‘span-FEM.NON.SG.’ (Thessalia, Tzartzanos 1909)
 (3d) /pu.ká.mi.so/ → [pká.msu] ‘shirt-NEUT.NOM.SG.’ (Meleniko, Andriotes 1989)
 (3e) /velonoθíci/ → [velun.θíci] ‘needle case-FEM.NOM.SG.’ (Meleniko, Andriotes 1989)
 (3f) /tu.fé.ci/ → [tfé.ci] ‘gun-NEUT.NOM.SG.’ (Thessalia, Tzartzanos 1909)

- (4a) /ku.bá.ros/ → [kba.ré.lɔ] ‘bestman-MASC.NOM.SG.’ (Thassos, Tombaidis 1967)
 (4b) /ku.vá.ri/ → [gvár] ‘ball-NEUT.NOM.SG.’ (Kozani, Roga 1989)*
 (4c) /ku.δú.ni/ → [kδu.nél] ‘bell-NEUT.NOM.SG.’ (Thassos, Tombaidis 1967)
 (4d) /tra.γy.dá.i/ → [tra.γdá.i] ‘sing-3SG.PRES.’ (Thessalia, Tzartzanos 1909)

⁷⁴ [+voi] + [-voi] combinations are subject to voicing assimilation (cf. Pater 1999).

⁷⁵ cf. Blaho & Bye (2006) for equivalent results.

(4e) /ti.yá.ni/ → [tyán] ‘frying pan-NEUTR.NOM.SG.’ (Samothraki, Katsanis 1996)

In most examples, except for (3d) and (3e), it is either both cluster members that are voiceless (3a-c, 3f,) or voiced (4a-b) or the leftmost member is a voiceless segment, whereas the rightmost is a voiced one (4c-e). Such data designate that there is a third scale, the voicing scale, which is responsible for cluster completeness. Like the scales of PoA and MoA, the voicing scale needs to be satisfied in a rightward direction, i.e. the first segment is voiceless and the second is voiced, and is vacuously satisfied in case both cluster members are either voiced or voiceless. This is illustrated in the data in (5). Given that the majority of clusters appearing in dialectal variants of Greek belong to the ‘acceptable clusters’ type, most clusters undergo (de)voicing assimilation, a common process cross-linguistically (cf. Al-Ahmadi Al-Habi to appear, Arvaniti 1999, Baroni 1997, Grijzenhout 2000).

Crucially, the voicing scale is violated if cluster members are selected in a leftward direction. In other words, leftmost voiced segments are prohibited, as displayed in (5e) and (5f), where the leftmost voiced segments undergo devoicing. The condition of rightward satisfaction of all scales is violated in case the leftmost segment is a nasal, as shown in (3d) and (3e) above. However, [nasal + voiceless obstruent] sequences are heterosyllabic, therefore, such cases are by definition excluded from the set of cases examined here. The voicing scale is depicted in figure 6 below.⁷⁶

- (5a) /pi.ðó/ → [bðó] ‘jump-1SG.PRES.’ (Thessalia, Tzartanos 1909)
 (5b) /pe.ði/ → [vði] ‘child-NEUT.NOM.SG.’ (Thessalia, Tzartanos 1909)
 (5c) /pa.ti.ma.sjá/ → [pa.tma.súð] ‘footmark-FEM.NOM.SG.’ (Thassos, Tombaidis 1967)
 (5d) /skou.dó/ → [gdó] ‘push-1SG.PRES.’ (Samothraki, Katsanis 1996)
 (5e) /po.di.kós/ → [pu.tkós] ‘mouse-MASC.NOM.SG.’ (Samothraki, Katsanis 1996)
 (5f) /ði.cé.li/ → [θcéli] ‘grub hoe-NEUT.NOM.SG.’ (Thassos, Tombaidis 1967)

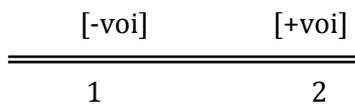


Figure 6: *The voicing scale*

It is important to note that the voicing scale is violated in one more case, i.e. in morpheme boundaries’ blending which is the product of vowel loss and results in the emergence of word final clusters. This is displayed in the data in (6). In these cases, the consonantal segments which make up the newly formed clusters retain their featural characteristics, consequently, the first segment is voiced and the second is voiceless. Again, these cases are excluded from the set of data examined here because the newly formed clusters are the product of surface processes rather than true phonological representations.

- (6a) /léjis/ → [léy+s#] ‘say-2nd.PRES.IND.SG.’
 (6b) /bakális/ → [bakál+s#] ‘grocer-MASC.NOM.SG.’ (Drimos, Katsanis 1983)

Table 4, like the equivalent tables in 1 and 2, summarizes cluster perfection and/ or (non)acceptability as well as gradience in cluster formation. A fundamental question underlying our theoretical proposal would be why to consider a distinct voicing scale and not assume the latter as being part of the classical sonority or MoA scale. A first potential answer would be that it is difficult to deal with CC cluster internal ‘voiceness’ without a distinct scale. More specifically, the data discussed in (3)-(6) exemplified that if the voicing scale is not satisfied clusters are not acceptable. As a result, and, in order to form

⁷⁶ Cf. also Grijzenhout & Kraemer (2000).

acceptable clusters, cluster segments undergo cluster assimilation. In addition, the data showed that the voicing scale is essentially one of the two scales that should be satisfied, given that violation of the conditions posed by the voicing scale is enough for clusters to be characterized as non-acceptable.

Table 4 : *Gradience in cluster formation (voicing)*

Types	Perfect	Accept	Non-accept
[-voi] + [-voi]		√	
[-voi] + [+voi]	√ /kδ/		
[+voi] + [+voi]		√ /gδ/	
[+voi] + [-voi]			√ /δk/

In section 2, we discussed gradience in cluster formation in detail. The discussion was summarized in table 3. In this section, we elaborate on table 3 by incorporating voicing in figure 7 below. In all boxes of figure 7, the leftmost cluster is the best of the category and the rightmost is the worst of this specific category. In the PC1 uppermost box, the most perfect among perfect clusters are provided. More specifically, we refer to clusters whose initial segment is voiceless and a second is (inherently) voiced. Given the above the most perfect among perfect clusters is /kl/ represented by [-voi]SV+L. It is important to remember that the most perfect cluster satisfies the place and manner scales; moreover, the distance among its members is the biggest possible, 4. The least perfect cluster, on the other hand, satisfies both scales of manner and place but the distance among its members is 1. Two of the least perfect clusters, /kθ/ and /kf/, appear in standard Greek but only in morpheme boundaries. In addition, /pθ/, the least perfect cluster, appears only in dialects. Similar hierarchies hold for acceptable and non acceptable clusters. AC1 represents acceptable clusters which are better formed compared to the leftmost AC2 clusters. Finally, non-acceptable clusters appear in the N-AC box. It is interesting to point out that two of the ‘worst’ non-acceptable clusters, /fk/ and /fx/, appear in standard Greek, though only in morpheme boundaries, as in ‘ef + kolos’ “easy” or ‘ef + xaristos’ “pleasant”.

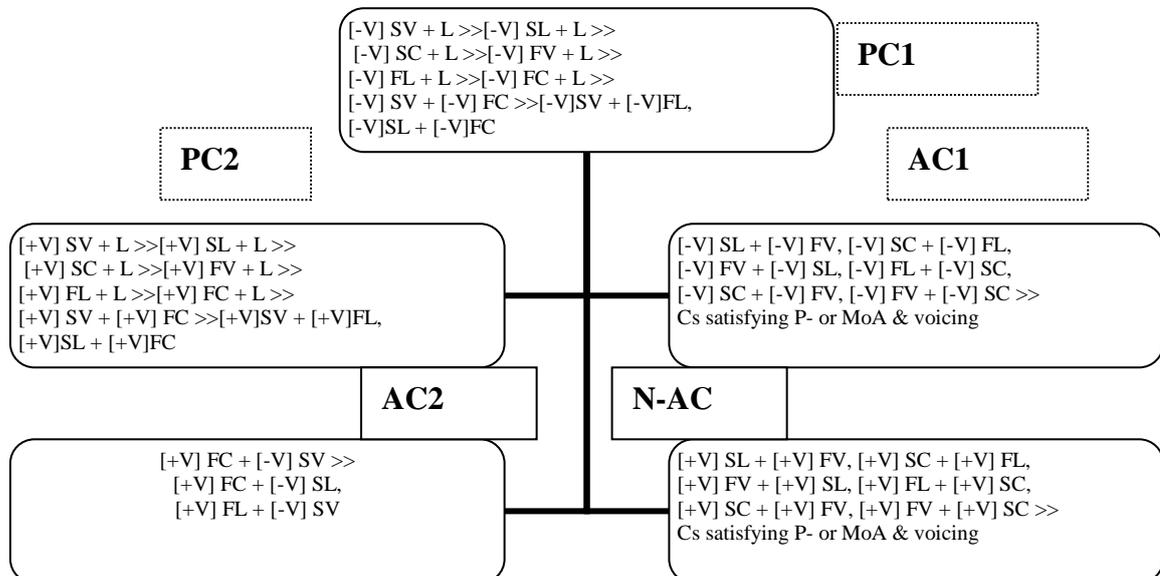


Figure 7: *gradiance in cluster perfection*

4. Concluding remarks and issues for future research

This paper assesses cluster well-formedness in a parallel fashion. More specifically, we evaluate cluster acceptability at the levels of, first, sonority, and, second, the distinctive features that determine segmental composition, namely, manner, place and voicing. Previous studies (Tzakosta & Karra in press) in combination with the present one have shown that clusters are divided in three categories of well-formedness, i.e. they are perfect, acceptable and non-acceptable. This categorization depends on the degree of satisfaction of the sonority scale as well as the scales of place and manner of articulation and voicing. Perfect clusters are sequences which respect the rightward direction of the sonority and the place, manner and voicing scales. Moreover, members of perfect clusters hold the biggest possible distance among them. To give an example, considering the sonority scale in figure 1, /pl/ and /pn/ are both perfect clusters, because they both respect the rightward direction of all scales. However, /pl/ is better-formed than /pn/ because /pl/ is characterized by sonority distance 5 while /pn/ is characterized by sonority distance 3. In other words, the bigger the distance among cluster members the better formed the cluster.

Acceptable clusters, on the other hand, are mainly CC clusters, i.e. sequences of segments highly adjacent on the sonority scale, like /pf/, or sequences of segments landing exactly on the same site on the sonority scale, like /fθ/. Adjacent segments, for example combinations of stops and fricatives, are maximally characterized by a sonority distance 1, while segments landing on the same sonority site, i.e. if both cluster members are stops or fricatives, are characterized by zero sonority distance. Consequently, acceptable clusters are characterized by coherence among their members. Put differently, acceptable clusters may violate one of the scales of place and/ or manner or vacuously satisfy one or both of these scales. It is interesting that, although place and manner may not be essentially or necessarily satisfied, voicing completes acceptable and/ or perfect cluster formation; therefore, it always needs to be satisfied. If the voicing scale is violated, the emergent cluster is non-acceptable.

The data reveal that, in theory, coherence is crucial for cluster survival, although, perception-wise, coherent – acceptable- clusters are not ‘true’ clusters (cf. Tzakosta & Vis 2009). Apparently, cluster coherence is responsible for the fact that acceptable clusters are the most frequent patterns which, in turn, drives the prediction that the latter are also dominant cross-linguistically.

Finally, non-acceptable clusters are consonantal sequences which violate the sonority scale and/ or one of the scales of place/ manner and voicing, i.e. its members are selected on a leftwards rather than a rightwards fashion. Non-acceptable clusters are the fewest in theory, a fact verified empirically by the data.

A final summarizing point in the discussion is that cluster formation, in general, and cluster perfection, in particular, is gradual in the sense that not all perfect or acceptable clusters are perfect or acceptable to the same extent. We still need to investigate our prediction that clusters acceptable in theory, like /fp/, /θf/, or /θt/, but not attested in the data are expected to emerge. More data need to be tested and classified.

References

- Αλεξάκης Ν. (2000). Παρατηρήσεις στη Φωνητική, τη Μορφολογία και τη Σύνταξη του Κερκυραϊκού ιδιώματος, *Μελέτες για την Ελληνική Γλώσσα 25*: 19-30.
- Ανδριώτης Π. (1989). *Το γλωσσικό ιδίωμα του Μελένικου*. Δημοσιεύματα της Εταιρείας Μακεδονικών Σπουδών, Θεσσαλονίκη.
- Αραβαντινού Π. (1909). *Ηπειρώτικον Γλωσσάριον*, Εκδ. Ν. Καραβία, Αθήνα.
- Arvaniti A. (1999). Greek Voiced Stops: Prosody, Syllabification, Underlying Representations or Selection of the Optimal?, *Proceedings of the 3rd International Conference of Greek Linguistics*, Ellinika Grammata, Athens, 883-890.

- Baroni M. (1997). The Representation of Prefixed Forms in the Italian Lexicon: Evidence from the Distribution of Intervocalic [s] and [z] in Northern Italian, M.A. Thesis, UCLA, 1997.
- Blaho S. & Bye P. (2006). *Cryptosonorants and the misapplication of voicing assimilation in Bispsectual Phonology*, ROA-759.
- Borowski T. (2000). Word Faithfulness and the Direction of Assimilation, *The Linguistic Review* 17: 1-28.
- Γιαννουλέλλης Γ. (1983). Πλωμάρι Λέσβου, Εκδ. Σ.Δ. Βασιλόπουλος, Αθήνα.
- Grijzenhout J. (2000). Voicing and Devoicing in English, German and Dutch: Evidence for Domain-specific Identity Constraints. *Working Papers Theorie des Lexikons* 116, Heinrich-Heine-Universität Duesseldorf.
- Grijzenhout J. & Kraemer M. (2000). Final Devoicing and Voice Assimilation in Dutch Derivation and Cliticization. In Stiebels, B. and D. Wunderlich (eds.) *Lexicon in Focus. Studia Grammatica* 45, Akademie Verlag, Berlin, 55-82.
- Kappa I. (2001). An Optimality Theoretic Account of the (West)Cretan Dialect. In Ralli, A., B. D. Joseph and M. Janse (eds.) *Proceedings of the First International Conference of Modern Greek Dialects and Linguistic Theory*, University of Patras, 105-116.
- Καραπατάκης Κ. (1976/1979). Ραμπόζι, Τσέτλες και Μαντάνια, *Πρακτικά Γ' Συμποσίου Λαογραφίας του Βορειοελλαδικού Χώρου (Ήπειρος-Μακεδονία-Θράκη)*, Ίδρυμα Μελετών Χερσονήσου του Αίμου, Θεσσαλονίκη.
- Κατσάνης Ν. Α. (1990-1991). Λεξιλόγιο Πλωμαρίου (Νομός Λέσβου), *Ελληνική Διαλεκτολογία* 2, Αφοι Κυριακίδη, Θεσσαλονίκη, 27-130.
- Κατσάνης Ν.Α. (1996). *Το Γλωσσικό Ιδίωμα της Σαμοθράκης*, Θεσσαλονίκη.
- Κατσάνης Ν. Α. (1981). Διαλεκτικά ΙΙ: Η Εξέλιξη της Φωνολογικής Αντίθεσης /λ/ : /λλ/ στο Ιδίωμα Δρυμού Μελισσοχωρίου και στα Περιφερειακά Ιδιώματα της Ελληνικής. *Μελέτες για την Ελληνική Γλώσσα* 6: 17-50.
- Κατσάνης Ν. Α. (1983). Τελικά Σύμφωνα - Συμφωνικά Συμπλέγματα του Ιδιώματος Δρυμού: Απόπειρα Συστηματοποίησής τους, *Μελέτες για την Ελληνική Γλώσσα* 8: 33-43.
- Κοντοσόπουλος Ν. (1997). *Θέματα Κρητικής Διαλεκτολογίας, Αναδημοσίευση μελετών*, Βιβλιοεκδοτική Αναστασάκη, Αθήνα.
- Κοντοσόπουλος Ν. (2006). *Γλωσσικός Άτλας της Κρήτης*. Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο.
- Lombardi L (1995). Laryngeal Features and Privativity, *The Linguistic Review* 12: 35-59.
- Lombardi L (1999). Positional Faithfulness and Voicing Assimilation in Optimality Theory, *NLLT* 17: 267-302.
- Μαλικούτη-Drachman A. (1998), Παρατηρήσεις στα Δασέα Σύμφωνα της Αττικής, *Μελέτες για την Ελληνική Γλώσσα* 18: 304-316.
- Μαργαρίτη-Ρόγκα Μ. (1989), Όροι Υφαντικής και Ενδυμασίας Καταφυγίου (Νομός Κοζάνης), *Ελληνική Διαλεκτολογία* 2, Αφοι Κυριακίδη, Θεσσαλονίκη, 173-180
- Μαργαρίτη-Ρόγκα Μ. (1990-1991). Όροι Υφαντικής και Ενδυμασίας Μικρού Κεφαλόβρυσου (Νομός Τρικάλων), *Ελληνική Διαλεκτολογία* 2, Αφοι Κυριακίδη, Θεσσαλονίκη, 131-135.
- Μελανοφρύδης Π. (1987). *Η εν Πόντω Ελληνική Γλώσσα*. Αφοι Κυριακίδη, Θεσσαλονίκη.
- Μηνάς Κ. (2004). *Μελέτες Νεοελληνικής Διαλεκτολογίας*, Τυπωθήτω-Δάρδανος, Αθήνα.
- Ντίνας Κ. Δ. (1985). Η Κώφωση-Αποβολή στα Βόρεια Νεοελληνικά Ιδιώματα (BNEI) και στο Κουτσοβλάχικο Ιδίωμα της Σαμαρίνας (KBΣ). *Μελέτες για την Ελληνική Γλώσσα*, 10: 57-66.
- Οικονομίδης Η. (1958). *Γραμματική της Ελληνικής Διαλέκτου του Πόντου*, Λεξικογραφικών Δελτίων, Παρ. 1, Ακαδημία Αθηνών, Αθήνα.
- Παντελίδου Χ. (1929). Φωνητική των Νεοελληνικών Ιδιωμάτων Κύπρου, Δωδεκανήσου και Ικαρίας, Αναδημοσιευμένη ως παράρτημα στην *Ελληνική Διαλεκτολογία* 2, Αφοι Κυριακίδη, Θεσσαλονίκη.
- Oostendorp van M. (2004). *An Exception to Final Devoicing*, ROA-656.
- Oostendorp van M. (2006), Incomplete Devoicing in Formal Phonology, Ms. Meertens Institute.
- Pater J. (1999). Austronesian Nasal Substitution and other NC effects. In Kager, R., H. Van der Hulst & Zonneveld (eds.) *The Prosody-Morphology Interface*, C.U.P. Cambridge, 310-343.
- Σετάτος Μ. (1994-1995). Ποντιακές Ετυμολογίες, *Ελληνική Διαλεκτολογία* 4, Εκδ. Οίκος Αδελφών Κυριακίδη, Θεσσαλονίκη, 25-38.
- Tzakosta M. (2004). Multiple Parallel Grammars in the Acquisition of Stress in Greek L1. Ph.D. Dissertation, LOT Dissertation Series 93, University of Leiden.

*“The importance of being voiced”:
cluster formation in dialectal variants of Greek*

- Tzakosta M. (2009). Asymmetries in /s/ Cluster Production and their Implications for Language Learning and Language Teaching, In Tsaggalidis, A. (ed.) *Selected Papers from the 18th ISTAL*, Department of English, Aristotle University of Thessaloniki, 365-373.
- Tzakosta M. (in prep.), *Gradient Well-Formedness in Consonant Cluster Formation: the Case of CL vs. CC Sequences*, Ms. University of Crete.
- Tzakosta M. & Karra A. (in press). A Typological and Comparative Account of CL and CC Clusters in Greek Dialects, *Proceedings of the 3rd International Conference of Modern Greek Dialects and Linguistic Theory*, University of Cyprus, Nicosia.
- Tzakosta M. & Vis J. (2009), Asymmetries of consonant sequences in perception and production: affricates vs. /s/ clusters, In Tsaggalidis, A. (ed.) *Selected Papers from the 18th ISTAL*, Department of English, Aristotle University of Thessaloniki, 375-384.
- Τζαρτζάνου Α. (1989). Περί της Συγχρόνου Θεσσαλικής Διαλέκτου. Αναδημοσιευμένη ως Παράρτημα στην *Ελληνική Διαλεκτολογία 1*. Αφοι Κυριακίδη, Θεσσαλονίκη.
- Τομπαΐδης Δ. (1967). Το Γλωσσικό Ιδίωμα της Θάσου, Διδακτορική Διατριβή. Επιστημονική Επετηρίδα Φιλοσοφικής Σχολής, Θεσσαλονίκη.
- Χατζιωάννου Κ. (1992/1993). Μερικές Αρχαΐζουσες Λέξεις της Κυπριακής Αθησαύριστες στα Εκδεδομένα Γλωσσάρια», *Κυπριακαί Σπουδαί ΝΣΤ'*: 11-17.
- Χατζιωάννου Κ. (1999). *Γραμματική της Ομιλούμενης Κυπριακής Διαλέκτου*, Με Ετυμολογικό Προσάρτημα, εκδ. Ταμασός, Λευκωσία.

Αντιγεγονοτικές περιφράσεις με το *ήθελα*+απρφ. στις νεοελληνικές διαλέκτους

ΣΥΜΕΩΝ ΤΣΟΛΑΚΙΔΗΣ
Πανεπιστήμιο Frederick
pre.st@frederick.ac.cy

1. Εισαγωγικά

Αντικείμενο της ανακοίνωσης είναι οι νεοελληνικές διαλεκτικές περιφράσεις με το *ήθελα*+απρφ. και αντιγεγονοτική λειτουργία. Στο πρώτο μέρος θα ασχοληθούμε κυρίως με την παρουσίαση των σχετικών δεδομένων, ενώ στο δεύτερο θα αναφερθούμε σε κάποια συμπεράσματα όσον αφορά: α) διάφορες γραμματικοποιητικές διαδικασίες που εμπλέκονται κατά το σχηματισμό των υπό εξέταση δομών, και β) κοινωνιογλωσσολογικού τύπου παράγοντες, όπως η επίδραση της κοινής νεοελληνικής σε νεοελληνικές διαλέκτους.

2. Τα δεδομένα

Οι πρώτες σποραδικές μαρτυρίες αντιγεγονοτικών δομών με το *ήθελα* ανάγονται στην ελληνιστική εποχή¹, π.χ.

1 (2ος αι. μ.Χ. ;) εἰ μὴ Ἰακώβ ὁ πατήρ ἡμῶν προσήυξατο περὶ ἐμοῦ πρὸς Κύριον, ἤθελε Κύριος ἀνελεῖν με "Ἄν ὁ Ἰακώβ ὁ πατέρας μας δεν προσευχόταν στον Κύριο για μένα, ὁ Κύριος θα με σκότωνε" (*Διαθήκαι των ΙΒ Πατριαρχῶν* 1040A).

αλλά δε φαίνεται να αρχίζουν να καθιερώνονται πριν ἀπὸ τὸ 14ο αἰ.², π.χ.

2 (14ος αἰ.) νὰ εἶχα ἐκεῖ εὐρεθῆ, ὅταν ἀπεπεζεῦθη ἀπάνω σου διὰ τὰ ἄρματα! Κακὰ τὰ ἤθελε πάρει "ἄν εἶχα βρεθεῖ ἐκεῖ, ὅταν ξεπέζεψε πάνω σου για να σου πάρει την πανοπλία! θα το εἶχε πληρώσει ακριβὰ" (*Πόλεμος της Τρωάδος* 4278-4279),

3 (15ος αἰ.) Ἄν τὸ ξεῦραν οἱ τεχνίτες, ἔθέλαν στοιχειῶνναι τοὺς ἐχθροὺς τους "Ἄν το ἤξεραν οἱ τεχνίτες, θα κατατρόμαζαν τους εχθρούς τους" (*Χρονικό Μαχαιρά* 592.14-15),

4 (16ος αἰ.) κλαύσειν ἤθελες, ἔὰν ὠργιζόμεον "θα ἐκλαιγες, ἄν ὀργιζόμεον" (*Πλουτάρχου Παιδαγωγός* 9),

5 (17ος αἰ.) πῶς ἔθέλαμε παχύνει "πῶς θα παχαίναμε" (*Κατζούρμπος* 3.135).

Στα (3) και (5) βλέπουμε ὅτι σε διαλεκτικό ἐπίπεδο ἡ συγκεκριμένη δομὴ μαρτυρεῖται στην **Κύπρο** και στην **Κρήτη**.

Στο *ήθελα*+απρφ. συνέβη επέκταση της συμφωνίας ως προς το πρόσωπο και τον αριθμό (agreement spreading)³, π.χ.

6 (17^{ος} αἰ.) ἀνίσως κι ἐκουδούνιζε, δεῖς ἤθελες "ἄν κουδούνιζε, θα ἐβλεπες" (*Κατζούρμπος* 1.197)

και στη συνέχεια περιορισμός της πλεονάζουσας και ἀπὸ τα δύο συστατικά της περίφρασης δήλωσης⁴.

Ἔτσι, θα μπορούσαμε να θεωρήσουμε ὅτι ἔχουμε μια αντιγεγονοτική δομὴ με πιο γραμματικοποιημένο τὸ α' συστατικό, με την ἔννοια ὅτι αὐτὸ με τὸ να καταστῆ ἀκλιτο, ἔχασε τα ρηματικά του χαρακτηριστικά (δήλωση προσώπου και αριθμοῦ), τα ὁποῖα τῶρα

¹ Ψάλτης (1918: 51)

² Markopoulos (2005: 205)

³ Βλ., ὁμως, και Markopoulos (2005: 214), που θεωρεῖ ὅτι ἡ δομὴ αὐτὴ σχετίζεται με σειριακὰ ρήματα (serial verbs).

⁴ Markopoulos (2005: 215). Σχετικὰ με τὸ ὅτι ὁ περιορισμός της πλεονάζουσας δήλωσης ἔγινε με βάση τὸ γ' εν., ἔχουμε ἐδῶ εφαρμογὴ του νόμου του Watkins, σύμφωνα με τον ὁποῖο τὸ γ' πρόσ. κατέχει τὴν πιο ουσιασθὴ θέση ὅσον ἀφορὰ τὴν ιστορικὴ εξέλιξη των ρηματικῶν παραδειγμάτων (βλ. Joseph 1980: 182).

δηλώνονται από το β' συστατικό της περίφρασης, που αποτελεί και το κύριο της ρήμα⁵. Σε επίπεδο νεοελληνικών διαλέκτων η δομή μαρτυρείται στις

7

Κεφαλονιά

1 *ithele perasune kala, ama itane pulotero aryatiki* "θα περνούσαν καλά, αν ήταν πιο εργατικοί"⁶,

Μύκονο

2 *an ithele mi soso na tone forto'ho* "μακάρι να μην προλάβαινα να τον φορτωθώ"⁷,

Πάτμο

3 *an ithele 'zisome* "αν ζούσαμε"⁸,

και

Κρήτη

4 *na thele se skotosune* "μακάρι να σε σκοτώνανε"⁹.

Εξάλλου, σύμφωνα με τον Pappas (2001: 87-88) περίπου στα μέσα του 16ου αι. εμφανίζεται η δομή *ήθελε*+πρω. κατ'αναλογία προς ανάλογες δομές μέλλοντα όπου το *θέλω* συνδυαζόταν με παρεμφατικό τύπο. Μια άλλη αιτία για την εμφάνιση μιας τέτοιας δομής θα μπορούσε να είναι ότι το άκλιτο *ήθελε* εξελίχθηκε σε έναν άκλιτο δείκτη τροπικότητας, που στη συνέχεια άρχισε να συντάσσεται με παρατατικό. Η εξέλιξη αυτή εντάσσεται στα πλαίσια της γενικής τάσης της μεσαιωνικής και νέας ελληνικής να σχηματίζει αντιγεγονοτικές δομές με παρατατικό¹⁰. Στη συνέχεια, μέσω επανερμηνείας του α' συστατικού του γ' εν., η συμφωνία ως προς το πρόσωπο και τον αριθμό επεκτάθηκε και στα υπόλοιπα πρόσωπα, με αποτέλεσμα να προκύψει το *ήθελα*+πρω., που μαρτυρείται στη

Σύρο

8 *a thelan kanan i skatobabuli meli* "αν μπορούσαν να κάνουν μέλι οι μπάμπουρες"¹¹.

Με βάση το *ίθελεν* και κατ'αναλογία προς τα ενεστωτικά γ' εν. *θε* "θέλει" και γ' πληθ. *θεν* "θέλουν" σχηματίστηκε το *ίθεν*¹², που μαρτυρείται στην

Πάτμο

9 *'ithe barome*¹³.

Θεωρούμε ιδιαίτερα αξιοσημείωτο το:

Μύκονος

10 *iθes erθi, ithele na su dosi mila* "αν ερχόσουν, θα σου έδινε μήλα"¹⁴,

⁵ Σύμφωνα με τον Lehmann (2002: 29-30) σε ένα υπό γραμματικοποίηση συνδυασμό λέξεων που αποτελείται από δύο ρηματικούς τύπους από τους οποίους ο ένας πρόκειται να εξελιχθεί στο βοηθητικό συστατικό μιας περιφραστικής δομής, αρχικά το βοηθητικό θα είναι αυτό που θα κυβερνά αλλά, στη συνέχεια, θα χάσει σταδιακά κάποια από τα ρηματικά του χαρακτηριστικά και, όταν πια θα έχει εξελιχθεί σε δείκτη τροπικότητας, αυτό που θα κυβερνά θα είναι το ρήμα που δηλώνει τη λεξική σημασία της περίφρασης. Επιπλέον, οι Horroger & Traugott (1993: 103-104) σημειώνουν ότι οι υπό γραμματικοποίηση τύποι τείνουν να χάνουν τα μορφολογικές ιδιότητες που τους χαρακτηρίζουν ως πλήρη μέλη μιας μείζονος γραμματικής κατηγορίας (όνομα ή ρήμα).

⁶ Κουστουράκης (1990: 43)

⁷ Μάνεσης (1997: 37)

⁸ Παπαδοπούλου (2004-2005: 178)

⁹ Markopoulos (2006: 240)

¹⁰ Βλ. Ψάλτη (1918: 44-45) και Horrocks (2006: 438).

¹¹ Μάνεσης (1997: 37)

¹² Psichari (1884: 25-27), Markopoulos (2005: 210)

¹³ Παπαδοπούλου (2004-2005: 178)

¹⁴ Μάνεσης (1997: 438). Όσον αφορά την έλλειψη του υποθετικού συνδέσμου, βλ. Horroger & Traugott (1993: 173) σχετικά με τον ρόλο του επιτονισμού για τη δήλωση μορφοσυντακτικών σχέσεων σε υποθετικές προτάσεις χωρίς υποθετικό σύνδεσμο. Το ίδιο φαινόμενο παρατηρείται και σε ορισμένες ποικιλίες της αγγλικής καθώς και σε υποθετικούς λόγους που σχηματίζονται με άλλες αντιγεγονοτικές δομές σε άλλες διαλέκτους, π.χ. στα Δωδεκάνησα

Χάλκη

επειδή το α' συστατικό της περίφρασης είναι κλιτό και ουσιαστικά μας παρέχει μια μαρτυρία για το ενδιάμεσο στάδιο προς εκείνη τη μορφή της περίφρασης με το α' συστατικό άκλιτο. Θυμίζουμε ότι στις μέχρι τώρα περιγραφές συνήθως γίνεται λόγος για το γ' εν. και γ' πληθ. του *ήθελα* που μέσω αναλογίας ή φωνολογικής μείωσης εμφανίζονται με τη μορφή *ήθεν*, αλλά σε όλες αυτές τις μαρτυρίες από τα πρωιμότερα κείμενα υπάρχει το πρόβλημα ότι δεν είναι ξεκάθαρο αν πρόκειται για κλιτό ή άκλιτο τύπου, π.χ.

11 (17^{ος} αι.) *νά'θεν εἶν' ἄλλοι τόσοι "να'ταν ἄλλοι τόσοι"* (Διήγησις Διγενή 2666).

Τα μυκονιάτικα μας παρέχουν σαφή – και τουλάχιστον απ'όσο γνωρίζω τη μοναδική – μαρτυρία για την περίφραση με το συντομευμένο τύπο να κλίνεται.

Μεγαλύτερο βαθμό γραμματικοποίησης έχουμε σε περιπτώσεις φωνολογικής μείωσης του *ήθεν*. Η μείωση αφορά είτε: α) το [i], είτε (β) το [θ].

Το (α) εντοπίζεται στην περιοχή των Κυκλάδων και παρατηρείται κατά τη σύμφυση του *iθe(n)* με το *an* ή το *as*. Στην πρώτη περίπτωση έχουμε το *aθθe(n)<anθe(n)* (με αφομοίωση [nθ>θθ])<*aniθe(n)*¹⁵ στις

12

Κίμωλο

1 *aθθe vrekxi, iθa jini* "αν έβρεχε, θα γινόταν (αυτό)"¹⁶,

και

Σίφνο

2 *aθθen pas na ton vris, iθa pines ce si gala* "αν πήγαινες να τον βρεις, θα έπινες κι εσύ γάλα"¹⁷,

ενώ με αφομοίωση [ae>ee] προέκυψε το *eθθe* στο **Καστελλόριζο** και με απλοποίηση [θθ>θ] το *aθen* στη

13

Μύκονο

1 *aθe bar fotja to stroma, simasia den idina. Tosi zest ixa* "φωτιά να'παιρνε το στρώμα, σημασία δε θα'δινά. Τόσο ζεστός ήμουν",

2 *aθen exo aletro* "μακάρι να είχα αλέτρι",

3 *aθen erθi, kala tane* "αν ερχόταν, θα ήταν καλά",

4 *aθe bas esi* "ας πήγαινες εσύ"¹⁸

Κατά τη σύμφυση του *iθen* με το *as* έχουμε το *aste<asθe* (με ανομοίωση ως προς τον τρόπο άρθρωσης)<*asiθe*¹⁹ στην

14

Κίμωλο

1 *aste mazoksi sporo* "ας μάζευε σπόρο"

2 *aste ercis prokthes, den iθena me mesazmeni* "αν ερχόσουν προχθές, δε θα'χα γίνει κομμάτια"²⁰.

Το ενδιάμεσο στάδιο *asθe* μαρτυρείται στη

Σίφνο

15 *makari asθe bao* "μακάρι να πήγαινα"²¹,

όπου, μάλιστα, φαίνεται ότι το *asθe* έχει καταστεί αδιαφανές σε τέτοιο βαθμό ώστε να επανενισχύεται με την προσθήκη του *μακάρι*. Το ίδιο ισχύει και για το 16.3 παρακάτω.

'icem 'pais e'si na ton vu'llosis ta 'stoma'ta to "ας πήγαινες εσύ να τους κλείσει τα στόματα" (Τσοπανάκης 1949: 64),

Πάτμος

'icem 'erti se 'mas na tone 'dume "ας είχε έρθει σε μας να τον δούμε" (Καραναστάσης 1956: 216).

¹⁵ Πβ. και Μάνεση (1997: 37) σχετικά με την αποβολή του μεσοσυμφωνικού [i].

¹⁶ Βογιατζίδης (1925: 157)

¹⁷ Ψάλτης (1918: 53-54)

¹⁸ Μάνεσης (1997: xxix, 37)

¹⁹ Βλ. και ΙΑ, λήμμα *άστε*.

²⁰ Βογιατζίδης (1925: 157)

²¹ Markopoulos (2006: 240)

Φυσικά, τα αποτελέσματα της σύμφυσης των *αν* και *ας* με το *ήθε* με το να γίνονται ολόένα και περισσότερο αντιληπτά ως ανεξάρτητοι (σε σχέση με το β' συστατικό της περίφρασης του *ήθε*) δείκτες τροπικότητας, άρχισαν να συνδυάζονται και με παρατατικό στις

16

Μύκονο

1 *aθe ganan* "αν κάναν",

2 *aθe do ksera na min erθo* "ας το'ξερα, να μην έρθω",

3 *maγar aθe gamune alo ena ksenodozio, na volevete o kosmos* «μακάρι να χτίζανε άλλο ένα ξενοδοχείο, να βολεύεται ο κόσμος»²²

και

Σίφνο

3 *asten irxusane noris, iθα me vris sto spiti mu* "αν ερχόσουν νωρίς, θα με έβρισκες στο σπίτι μου"²³.

Η φωνολογική μείωση τύπου (β) έδωσε στην Κύπρο το *ien*²⁴, π.χ.²⁵

17

1 *an ien to ksevro* "αν το'ξερα"

2 *an iel lipun* "αν έλειπαν",

3 *as ien erto xtes c esintixanamen* "ας ερχόμουν χθες και θα μιλούσαμε".

Με αποβολή του αρχικού άτονου φωνήεντος των *anien* και *asien* προέκυψαν τα *nien*²⁶ και *sien* αντίστοιχα, π.χ.

18

1 *nien ertis, esazamen ta* "αν είχες έρθει, θα τα τακτοποιούσαμε",

2 *nien men erto* "μακάρι να μην ερχόμουν".

3 *para naen pais is tin horan, sien ertis da* "αντί να πήγαινες στη Χώρα, ας ερχόσουν εδώ".

Εξάλλου από τη σύμφυση του *ien* με το *na* προέκυψε το *naen*<*naien* με έκκρουση του [i] κοντά στο [a]²⁷, π.χ.

19

1 *para naen pais is tin horan, sien ertis da* "αντί να πήγαινες στη Χώρα, ας ερχόσουν εδώ",

2 *naen to ksero, en itan na rto* "αν το ήξερα, δε θα ερχόμουν",

3 *naen kai i ora pu ton epandreftika* "καταραμμένη η ώρα που τον παντρεύτηκα".

Επιπλέον, με στένωση [ae]>[ai]²⁸ το *naen* εξελίχθηκε σε *nain*, π.χ.

20

1 *nain kai i ora pu ton epandreftika* "καταραμμένη η ώρα που τον παντρεύτηκα",

ενώ κατ'αναλογία προς το *naen* έχουμε το *saen*²⁹, π.χ.

²² Μάνεσης (1997: xxix, 37)

²³ Ψάλτης (1918: 54)

²⁴ Όσον αφορά το *αν* το *ien* προέρχεται από το *είχεν* ή από το *ήθεν*, βλ. τη σχετική συζήτηση στον Μηνά (1975: 135-140). Με βάση και την παρατήρηση του Markopoulos (2005: 220) ότι βασικό διαφοροποιητικό χαρακτηριστικό των περιφράσεων με το ΘΕΛΩ είναι πως μόνο αυτές υφίστανται φωνολογική μείωση, τώρα πια θεωρούμε απλώς ως πιθανότερη την προέλευση από το *ήθεν*. Βέβαια, και όσον αφορά το ΕΧΩ, στη Β. Ελλάδα μαρτυρείται το *xana*<*ixana* (Τζιτζιλής, υπό έκδ.), οπότε σε αυτή την περίπτωση θα μπορούσαμε να θεωρήσουμε ότι και στην Κύπρο το *ien* θα μπορούσε να έχει προκύψει από το *ίzen* με αποβολή του [ç]. Το πρόβλημα έγκειται και στο γεγονός πως λόγω του ότι το *ien* ανήκει περισσότερο στον προφορικό κώδικα, ακόμα και αν είχε εμφανιστεί πριν από το *ίθεν* (οπότε τότε θα προερχόταν αναγκαστικά από το *ίzen*), δεν είναι εύκολο να έχουμε τέτοιες μαρτυρίες. Σε αυτή την περίπτωση τίποτε, βέβαια, δεν αποκλείει το *ίθεν* να αντικατέστησε ή να συνυπήρξε με το παλαιότερο *ίzen* ή τις οποιεσδήποτε εξελίξεις του και στη συνέχεια μέσω φωνολογικής μείωσης να εξελίχθηκε και αυτό σε *ien*, κάτι που σημαίνει ότι το *ien* θα μπορούσε να προέρχεται και από το *ίzen* και από το *ien*.

²⁵ Μενάρδος (1925: 46), Χατζηγιάννου (1999:94).

²⁶ Μενάρδος (1925: 47)

²⁷ Μηνάς (1975: 137). Βλ. και Μενάρδο (1925: 47) και Χατζηγιάννου (1999: 93).

²⁸ Βλ. τη σχετική συζήτηση για τη στένωση στον Μωυσιάδη (2005: 88, 95-96).

²⁹ Μενάρδος (1925: 48). Για μια διαφορετική ετυμολόγηση βλ. Χατζηγιάννου (1999: 94).

2 *saen ezisi c as ta xannamen* "ας ζούσε κι ας τα χάναμε",
 3 *saen pai* "ας είχε πάει",
 που με στένωση [ae]>[ai] εξελίχθηκε σε *sain*, π.χ.
 4 *sain pai* "ας είχε πάει".

Τέλος, σύμφωνα με την περιγραφή του Μενάρδου, τόσο το *nien* όσο και το *naen* σε ένα συγκεκριμένο περιβάλλον συνδυάζονται με παρατατικό, π.χ.

21

1 *sjan nien itaδ dikon tu*,
 2 *sjon naen ituδ dikon tu* "σα να'ταν δικό του".

κάτι που ίσως να οφείλεται στην επιθυμία των ομιλητών για αυξημένη εκφραστικότητα³⁰, με την έννοια ότι τα *nien* και *naen*, τα οποία είχαν εξελιχθεί σε ανεξάρτητα τροπικά μόρια, εισήχθησαν σε ένα περιβάλλον όπου χρησιμοποιούταν απλός παρατατικός, προκειμένου τονιστεί ακόμα περισσότερο η αντιγεγονοτικότητα³¹.

Τέλος, όσον αφορά το *asen* της **Αστυπάλαιας**³², π.χ.

22

1 *asen erti ts emen to pedim mu po ti ksenedtza* "μακάρι να'ρχόταν το παιδί μου από την ξενιτιά",

2 *asen to xero peritsi pos theta s exo feti* "ας το'ξερα από πέρυσι ότι θα σ'έχω φέτος"³³,
 θεωρούμε ότι μπορεί να προέκυψε από το *asien*(*asiθen*) με έκκρουση του [i] κοντά στο [e]³⁴.

3. Συζήτηση-συμπεράσματα

Τα δεδομένα που παρουσιάσαμε δείχνουν ότι η ελληνιστική ή μεσαιωνική ελληνική περίφραση *ήθελα*+απρφ. χρησιμοποιήθηκε σε διάφορες περιοχές του νησιωτικού ελληνόφωνου χώρου (Κυκλάδες, Επτάνησα, Κρήτη, Δωδεκάνησα, Κύπρος)³⁵ με διαφορετικούς βαθμούς ή διαδικασίες γραμματικοποίησης που συμπεριλαμβάνουν αλλαγές στη φωνολογία, στη μορφολογία, στη σύνταξη και στη σημασιολογία, κάτι που και σε επίπεδο ΝΕ διαλέκτων δικαιώνει το χαρακτηρισμό της γραμματικοποίησης ως μιας διαεπιπεδικής (cross-componential) μεταβολής³⁶.

Έτσι, τα παρ. 9-22 θα μπορούσαν να θεωρηθούν περιπτώσεις απώλειας παραδειγματικού βάρους (paradigmatic weight) με φωνολογική μείωση (phonological attrition)³⁷, η οποία, όπως βλέπουμε στα παραδείγματα (12-22) συνοδεύεται από αύξηση της δεσμευτικότητας (bondedness) και σαφή συγχώνευση (coalescence) με το γειτονικό φωνητικό υλικό. Οι Bybee, Perkins & Pagliuca (1994: 6-7, 110) σημειώνουν ότι όσο αυξάνεται η μείωση σε επίπεδο φωνητικής και σημασίας, τόσο πιο πολύ αυξάνει η τάση για συγχώνευση με το περιβάλλον φωνητικό υλικό, ενώ σύμφωνα με τον Lehmann (2002: 132-133) θα μπορούσαμε να θεωρήσουμε ως προϋπόθεση για τη σύμφυση τη

³⁰ Σχετικά τη σχέση της γραμματικοποίησης με την επιθυμία για αυξημένη εκφραστικότητα βλ. Meillet (1912: 139-140) και Lehmann (1985: 315).

³¹ Η εμφάνιση του *nien* και *naen* στο συγκεκριμένο περιβάλλον πιθανότατα διευκολύνθηκε και λόγω της τάσης που, σύμφωνα με τη διαγλωσσική έρευνα των Bybee, Perkins & Pagliuca (1994: 26) έχουν οι δείκτες τροπικότητας να εμφανίζονται σε περιβάλλοντα δευτερευουσών προτάσεων με τα οποία ταιριάζουν σημασιολογικά και όπου η σημασιολογική συνεισφορά των δεικτών δεν είναι ανεξάρτητη [«where their meaning harmonizes with the (ενν. subordinate) context instead of making an independent contribution»].

³² Καραναστάσης (1958: 132)

³³ Καραναστάσης (1958: 132)

³⁴ Βλ. τη σχετική συζήτηση για την έκκρουση στον Μωυσιάδη (2005: 85-86)

³⁵ Σε καμιά περίπτωση δεν υπαινισσόμαστε εδώ την ύπαρξη κάποιου ισόγλωσσου, για τον λόγο ότι, όπως είναι γνωστό, τα δεδομένα που διαθέτουμε για την κατανομή διαφόρων γλωσσικών φαινομένων των νεοελληνικών διαλέκτων (ένα από τα οποία είναι και οι δυνητικές περιφραστικές δομές) δεν είναι ακόμα πλήρη.

³⁶ McMahon (2003: 232)

³⁷ Βλ. Lehmann (2002: 113).

σημασιοσυντακτική σχέση (grammatical relation) των δύο στοιχείων που θα συμπεροφερθούν. Και η McMahon (2003: 245-248) σημειώνει ότι τα μορφολογικά στοιχεία που βρίσκονται το ένα κοντά στο άλλο, είναι αυτά που έχουν και μια πολύ στενή σημασιολογική σχέση, ενώ, στη συνέχεια στα πλαίσια της γραμματικοποίησης ακολουθεί η συγχώνευσή τους. Η σημασιολογική εγγύτητα είναι απαραίτητη για τη συγχώνευση, αφού το αποτέλεσμα της θα πρέπει να γίνεται αντιληπτό ως μια ενιαία σημασιολογικά ερμηνεύσιμη μονάδα, όπως συμβαίνει, π.χ., στην περίπτωση των *aθθεν* και *asten*, τα οποία προήλθαν από μια συντακτική δομή όπου το δυνητικό μόριο ή σύνδεσμος και η αντιγεγονοτική περίφραση συνυπάρχουν δίπλα δίπλα σε προτάσεις που εισάγονται με το σύνδεσμο ή το μόριο και εκφέρονται με τη δυνητική περίφραση. Βλ. π.χ. στο 7.2 *an iθele mi soso na tone forto 'θo* "μακάρι να μην προλάβαινα να τον φορτωθώ" το *an iθele*+υποτ., τον πιθανότατο, σύμφωνα με τον Μάνεση (1997: 37), πρόγονο του *aθε* (βλ. παρ. 16.1 *aθε ganan* "αν κάναν") στην περίπτωση της Μυκόνου.

Μια άλλη παρατήρηση που θα μπορούσαμε να κάνουμε όσον αφορά τα *aθθεν*, *asten*, *nien* και *naen*, είναι ότι και σε νεοελληνικές διαλέκτους βλέπουμε να εμφανίζονται υποθετικοί σύνδεσμοι που προέρχονται από τροπικά στοιχεία. Ανάλογες παρατηρήσεις σε διαγλωσσικό επίπεδο κάνουν οι Hopper & Traugott (1993: 179).

Όσον αφορά την ποικιλομορφία στο εσωτερικό διαλέκτων, δηλ. στο εσωτερικό του ίδιου γλωσσικού συστήματος, είναι γνωστό ότι η εμφάνιση ενός καινούργιου γραμματικού δείκτη δεν εξαρτάται από την απώλεια των ήδη υπαρχόντων³⁸. Χαρακτηριστικά, βλέπουμε στη Μύκονο ότι, εκτός από τα παρ. 7.2 και 16.1, έχουμε και το 10 *iθes erθi, iθele na su dosi mila "αν ερχόσουν, θα σου έδινε μήλα"*, όπου στον ίδιο υποθετικό λόγο έχουμε στη μεν υπόθεση τη δομή που προέρχεται από το *ήθελα*+απρφ., στη δε υπόθεση τη δομή που ανάγεται στο *ήθελα να*+παρεμφατικός τύπος³⁹. Μάλιστα, η συγκεκριμένη κατανομή των δομών αυτών στο παρ. 10 έχει να κάνει με το γεγονός ότι η παλαιότερες δομές εμφανίζονται σε πιο εξειδικευμένα περιβάλλοντα (όπως είναι οι υποθετικές προτάσεις). Στη Μύκονο οι δομές που ανάγονται στο *ήθελα να*, δεν εμφανίζονται σε υποθετικές προτάσεις, όπου εκτός από την περίφραση με το απρφ., μπορεί να εμφανιστεί και απλός παρατατικός (άλλη μια δομή σαφώς παλαιότερη από την *ήθελα να*+παρεμφατικός τύπος), π.χ.

23 *a den eruvazizes, iθele na peθano "αν δεν ερχόσουν, θα πέθαινα"*⁴⁰.

Τέλος, όσον αφορά τις αντιγεγονοτικές δομές με παρατατικό, θεωρούμε ότι μπορεί να σχετίζονται άλλοτε με εσωτερικούς παράγοντες (π.χ. την τάση της ελληνικής να χρησιμοποιεί τον παρατατικό με αντιγεγονοτική λειτουργία⁴¹) και άλλοτε σε επίδραση αντιγεγονοτικών δομών της κοινής νεοελληνικής με παρατατικό, π.χ. στην περίπτωση της Μυκόνου θεωρούμε ότι στο

13.4 *aθε bas esi "ας πήγαινες εσύ"*

έχουμε μια δομή παλαιότερη από αυτήν του 16.2, η οποία μπορεί να οφείλεται στην επίδραση αντίστοιχων δομικών σχημάτων της κοινής νεοελληνικής με παρατατικό. Σε άλλες περιπτώσεις, μπορεί η συνύπαρξη διαφορετικών δεικτών να έχει να κάνει με επίδραση όχι της κοινής νεοελληνικής αλλά άλλων διαλεκτικών συστημάτων, π.χ. στην περίπτωση της Κιμώλου βλέπουμε ότι σε υποθετικές προτάσεις χρησιμοποιείται και το *aθθε* (βλ. παρ. 12.1) και το *aste* (βλ. παρ. 14.2). Τίποτα δεν αποκλείει το ένα από τα δύο να οφείλεται σε επίδραση κάποιας άλλης κυκλαδίτικης ποικιλίας (π.χ. μέσω κάποιας μετακίνησης πληθυσμών). Γενικά, οι Hopper & Traugott (1993: 114) παρατηρούν ότι η χρήση περισσότερων του ενός σχηματισμών με την ίδια λειτουργία σε συγχρονικό επίπεδο μπορεί να σχετίζεται με κοινωνιογλωσσολογικούς παράγοντες και, βέβαια, δεν υπάρχει κανένας λόγος να αποκλίνουν από αυτό και οι ποικίλες νεοελληνικές

³⁸ Hopper & Traugott (1993: 125), Bybee, Perkins & Pagliuca (1994: 21-22)

³⁹ Βλ. Τσολακίδης (2009) για την εξέλιξη της συγκεκριμένης δομής στις νεοελληνικές διαλέκτους.

⁴⁰ Μάνεσης (1997: 348)

⁴¹ Βλ. Horrocks (1995). Σε διαγλωσσικό επίπεδο βλ. τις παρατηρήσεις των Bybee, Perkins & Pagliuca (1994: 233) για παρόμοια χρήση του παρατατικού στην αρμενική.

αντιγεγονοτικές δομές με το ήθελα+απρωφ. Τέλος, τίποτα δεν αποκλείει τη συνεπίδραση παραγόντων, π.χ. σε κάποια διάλεκτο η γενική τάση της ελληνικής για τη χρήση του παρατατικού σε αντιγεγονοτικές δομές να ενισχύθηκε από την επίδραση δομών της κοινής νεοελληνικής.

Βιβλιογραφία

- Ιστορικών λεξικόν της νέας ελληνικής της τε κοινώς ομιλουμένης και των ιδιωμάτων*, τ. 1-5β (1933-1989). Αθήνα, Ακαδημία Αθηνών.
- Βογιατζίδης Ι. (1925). Περί της γλώσσης των Κιμωλίων. *Αθηνά* 36, 109-160.
- Καραναστάσης Α. (1958). Το ιδίωμα της Αστυπάλαιας. *Λεξικογραφικόν Δελτίον* 8, 59-144.
- Κουστουράκης Γ. (1988). *Δήμος Άσσου Κεφαλληνίας*. Αθήνα: Αχαϊκές Εκδόσεις.
- Μάνουρας Σ. (1997). *Ιστορικό λεξικό του μυκονιάτικου ιδιώματος*. Μύκονος: Δήμος Μυκονίων.
- Μενάρδος Σ. (1925). Κυπριακή γραμματική Γ'. *Αθηνά* 37, 35-79.
- Μηνάς Κ. (1975). Γλωσσικά ανάλεκτα. *Δωδώνη* 4, 135-145.
- Μωυσιάδης Θ. (2005). *Ετυμολογία*. Αθήνα: Ελληνικά Γράμματα
- Παπαδοπούλου Μ. (2004-2005). Το ιδίωμα της Πάτμου. *Λεξικογραφικόν Δελτίον* 25, 169-196.
- Τζιτζιλής Χ. (επ.) (υπό έκδ.). *Οι νεοελληνικές διάλεκτοι*. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών.
- Τσολακίδης Σ. (2009). Αντιγεγονοτικές περιφράσεις με το ήθελα να στις νεοελληνικές διαλέκτους. *Μελέτες για την ελληνική γλώσσα. Πρακτικά της 29ης Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας του Α.Π.Θ. (Θεσσαλονίκη, 10-11/5/2008)*. Θεσσαλονίκη: Ινστιτούτο Νεοελληνικών Σπουδών, 413-427.
- Χατζηϊωάννου Κ. (1999). *Γραμματική της ομιλούμενης κυπριακής διαλέκτου*. Λευκωσία: Ταμασός.
- Ψάλτης Σ. (1918). Περί των υποθετικών λόγων εν τη μέση και νεωτέρα ελληνική. *Λεξικογραφικόν δελτίον* 5, 40-57.
- Bybee J., R. Perkins & W. Pagliuca (1994). *The evolution of grammar*. Chicago: University of Chicago Press.
- Heine B. & T. Kuteva (2002). *World lexicon of grammaticalization*. New York: Cambridge University Press.
- Horrocks G. (1995). On condition...: aspect and modality in the history of Greek. *Proceedings of the Cambridge Philological Society* 41, 153-173.
- Horrocks G. (2006). *Ελληνικά: ιστορία της γλώσσας και των ομιλητών της*. Αθήνα: Εστία.
- Hopper P. & E. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Joseph B. (1980). Watkins' Law and the Modern Greek Preterite. *Die Sprache* 26, 179-184.
- Lehmann C. (1985). Grammaticalization: synchronic variation and diachronic change. *Lingua e stile* 20.3, 303-318.
- Lehmann C. (2002). *Thoughts on grammaticalization*. Erfurt: Universität Erfurt.
- Markopoulos T. (2005). *The category 'Future' in Greek*. PhD diss. στο University of Cambridge.
- Markopoulos T. (2006). The development of future/modality markers: evidence from Modern Greek dialects. *Janse M., B. Joseph & A. Ralli (eds.), Proceedings of the 2nd International Conference of Modern Greek Dialects and Linguistic Theory (Mytilene, 30/9-3/10 2004)*. Patras: University of Patras, 236-255.
- McMahon A. (2003). *Ιστορική γλωσσολογία: η θεωρία της γλωσσικής μεταβολής*. Αθήνα, Μεταίχμιο.
- Meillet A. (1912). *Linguistique historique et linguistique générale*. Paris: Champion.
- Pappas P. (2001). The microcosm of a morphological change: variation in *thelō*+infinitive Futures and *ēthela*+infinitive counterfactuals in Early Modern Greek. *Diachronica* 18, 59-92.
- Psichari J. (1884). Futur composé du grec moderne. *Mémoires de la Société de Linguistique de Paris* 5, 3-47.

moro, akro, sjo: Prefixes or Compound Constituents?

Eleonora Dimela
University of Patras
eldimela@upatras.gr

Angela Ralli
University of Patras
ralli@upatras.gr

Abstract

In this paper we examine the items *akro-*, *moro-* and *sjo-* that constitute three particular cases of grammaticalization within the morphological domain. Using data from the Modern Greek dialects, we show that for an item to be a lexeme or a prefix depends on specific phonological, semantic and morphological properties. These properties indicate the specific parameters which are involved in a morphologization process.

1. Defining Prefixization

1.1 Grammaticalization vs. Morphologization

The classical concept of grammaticalization (or grammaticization, or even grammatization) originates from Meillet (1912: 131), who has defined it as “the attribution of a grammatical character to a previously autonomous word”. As noted by Hopper (1991: 18), grammaticalization for Meillet refers to an array of forms, which constitute the morphology of a language. It is only latter (following work by Givón 1971, 1979, Heine & Reh 1984, Lehmann 1985, Hopper & Traugott 1993, McMahon 1994, Gaeta 1998) that the range of grammaticalization phenomena are shown not to be restricted to morphology, and that the process is seen as encompassing all types of language change, having a broader scope as the study of the origins of grammar in general.

The emergence of elements with a morphological role from items which were not a matter of morphology in a previous stage is usually called ‘morphologization’. Hopper & Traugott (1993: 135) define morphologization as the creation of a bound morpheme out of an independent word, and Joseph (2003) discusses two types of morphologization, namely desyntactization and dephonologization, on the assumption that there is a wide range of phenomena that show ‘movement into morphology’ (see also Klausenburger 2002). Joseph argues that morphologization has to be kept distinct from grammaticalization, although the two may overlap to some extent: on the one hand, grammaticalization may make claims about changes that have nothing to do with morphology, and on the other hand, morphologization may involve changes that can be accommodated within morphology, but do not involve the grammar in general (see Joseph 2003 for more details).¹

In this presentation, we investigate a prefixization process in Greek, which is developed out of compounding. We examine a number of items which appear in morphologically complex words, and have become, or tend to become, prefixes. As Ralli (2007, 2009a, 2010) has shown, Greek compounding and prefixation are morphological processes according to the following basic criteria:

- a) Compound and prefixed formations display one stress, i.e. they are single phonological words,
- b) They involve bound elements. On the one hand, Greek prefixes are non-separable entities, and on the other hand, Greek compounding is mainly stem based², since, with

¹ For instance, Joseph (2003: 47) criticizes the formation of the German word *heute* ‘today’ from a presumed instrumental phrase **hiu tagu* in Old High German, since “...this combination of sounds is as grammatical (or not, as the case may be) before the phrase was reduced as it is afterwards”.

² According to Ralli (2005, 2009a, 2009b), in Modern Greek, there is no structural difference between a root and a stem, as opposed to Ancient Greek, where stems were usually combinations of

some exceptions, the first component is a stem, while the second component can be either a stem or a word.

c) They are subject to word atomicity, i.e. no syntactic rules have access to their internal structure.

d) In many cases, prefixed and compound formations are semantically non-compositional.

In this respect, moving from compounding to prefixation takes place within morphology, and should be seen as an instance of morphologization, in the sense that prefixation implies a greater morphological involvement than compounding, since stems display a greater autonomy than prefixes. For instance, stems can be used as independent words with the appropriate inflectional ending, and have a specific lexical meaning. In contrast, prefixes cannot be used as autonomous entities, and have a rather abstract semantic function, which contributes to the determination of the meaning of the word.³

It should be noticed that the difference between prefixes and compound constituents is accounted for by certain frameworks (see, among others, Anderson 1992, within the framework of a process-morphology model), which assign to compounds a rather syntactic structure, while they realize prefixes by morphological operations. Under this perspective, prefixation could also be seen as an instance of grammaticalization in the classical sense, that is as a process where lexemes acquire a grammatical role.

On the basis of the considerations above, one may argue that prefixation involves movement along a scale ('cline') of increasing grammatical status, by which expression via prefixation can be considered as 'more morphological' as well as 'more grammatical' than expression via compounding. Given the fact that prefixation deriving from compounding is an instance of morphologization, as well as of grammaticalization, we prefer adopting the term of morphologization, since, as also pointed out by Joseph (2003: 478), in grammaticalization studies there is a tendency to ignore the formal question of where in the grammar a particular phenomenon is located.

1.2 Parameters of prefixation

It is generally accepted that grammaticalization occurs if certain criteria are satisfied, which correspond to a number of parameters accounting for the process (see, among others the theoretical approaches of Lehmann [1982] 1995, Hopper 1991, Heine 2003, Heine & Kuteva 2002, 2005, 2007, Amiot 2005, Marchello-Nizia 2006, van Goethem 2007, 2008). As far as prefixation is concerned, and with some degree of variation from one author to another, there is more or less agreement on the following general criteria:⁴

- Phonological erosion
- De- or re-semanticization
- Decategorialization (or transcategorialization according to Ramat 2001)
- Extension

According to Joseph (2003: 477), each of these criteria is in principle independent of the others, and grouping them is purely stipulative. Moreover, to our knowledge, there are no explicit proposals about the order according to which these criteria are met for a prefixation process, with maybe the exception of Booij (2005), who claims that semantic change precedes the formal one.

With respect to these observations three basic questions are raised: a) what are the specific parameters which induce prefixation out of compounding? b) Are these parameters the same for all the range of grammaticalization phenomena, or are restricted

roots and thematic vowels. Today, the notion of a thematic vowel is no longer relevant. See Ralli 2007, 2009a, in preparation, for more details on Greek compounds.

³ See Iacobini (2004) for a range of abstract meanings, which may be assumed by a prefix.

⁴ Paradigmatization has also been proposed by Lehmann (1985) as a parameter for an item to be grammaticalized. This parameter refers particularly to inflection, which has a typical paradigmatic character, while for prefixes, this parameter is meaningful only if we consider them to be distributed into specific paradigms. See van Marle (1985) for the notion of paradigms in derivation.

to morphology, and thus to morphologization? c) Is there a particular order according to which these parameters occur?

Following the general assumption that morphology is an independent grammatical module with its own rules and constraints, our position is that if we deal with morphology the parameters which lead to the completion of a morphologization process have to be morphological in nature. Other parameters may trigger the process, or may play a role during the process, but do not guarantee completion. Within this spirit, we propose that the general grammaticalization parameters which may be involved in prefixization are resemanticization and erosion, but the specific morphological parameters, which are crucial for determining the final stage of prefixization, are related with

- the expansion of morphological combinatorial properties, and
- the increase of productivity.

We also suggest that the parameters playing a role in prefixization are not of equal weight. In an effort to rate their importance we show that:

a) Resemanticization is compulsory for an item to become a prefix (as also correctly noticed by Booij 2005), but does not guarantee completion.

b) Erosion⁵ (in accordance with Heine & Kuteva 2007) may play a role in prefixization, but it is not a necessary condition for the process to start, or to be achieved. Furthermore, it may precede semantic and morphological change.

c) The increase of productivity and the expansion of morphological combinatorial properties are strong indications for a candidate to have reached the final stage of morphologization (see also van Goethem 2008 and Amiot 2005 respectively).

Finally, we consider decategorialization to be the result of morphologization, but not part of the process itself.

These suggestions imply a certain degree of hierarchical application of the parameters involved in prefixization: Desemanticization and phonological erosion precede the morphologically-proper parameters, which, in their turn, lead to decategorialization.

Support for these proposals comes from research in the dialectal domain. We use evidence from several Modern Greek Dialects, where three particular items, *akro*, *moro* and *sjo*, originate from nominals, but have become, or tend to become prefixes, each one demonstrating a number of peculiar properties. The dialectal data are drawn from local dictionaries, grammars, dialectal documents, the archives of the Centre of Modern Greek Dialects of the Academy of Athens, and the oral material of the Laboratory of Modern Greek Dialects of the University of Patras.

1.3 Prefixation vs. Compounding in Greek

There is more or less agreement among linguists (cf., among others, Iacobini 2004, Stekauer 2005) that typical prefixes display the following properties: they are category neutral, occupy a particular position within prefixed words (preposed to a constituent), are structurally dependent on the base, and do not have a specific lexical meaning. Non-separability, or loss of lexical autonomy may be another property (Iacobini 2004, Booij 2005), but as shown by van Goethem (2007), separability is not a decisive criterion to define an item as a prefix.

As opposed to prefixes, items participating in Greek compounding bear a specific grammatical category (at least for languages like Greek, where there are no verbs and nouns sharing the same form⁶), may appear first or second elements in compound

⁵ We prefer using the term *erosion* than *phonological reduction*, since as pointed out by Heine & Kuteva (2007: 44), the former implies a wider sense and it can be linked to grammaticalization phenomena.

⁶ With the exception of a handful of stems (e.g. *kinig(os)* 'hunter' vs. *kinig(o)* 'to hunt'), which share the same form in both verbs and nouns, and only their inflectional endings are different. However, this is not sufficient evidence in order to adopt a model like that of *Distributed Morphology*, where

formations, may or may not be structurally dependent on the base (see subordinate vs. coordinate compounds), and have a specific lexical meaning (see Ralli 2009, 2010, in preparation).

However, between the two categories, prefixes and stems, there is no radical separation. There are items, the so-called ‘affixoids’ (Fleisher 1969), which may share properties with both categories: an increased productivity, a decreased semantic specificity, and a link to an existing free stem. As noted by ten Hacken (2000: 355), the first two criteria make affixoids resemble affixes, while the third one distinguishes affixes from affixoids. Following Ralli (2005, 2010), the intermediate category of affixoids can justify the existence of a morphological cline, where the two poles are occupied by typical affixes (prefixes and suffixes) and stems, while affixoids are situated in between. Moreover, the existence of affixoids can also motivate a cline of morphologization, which denotes the fact that the morphological change from compounding to prefixation is gradable,⁷ and that there are intermediate stages demonstrating that the boundaries between the two processes are not very clear (Booij 2005, Bauer 2005, Ralli 2010).

2 The data

2.1 *akro-*

In Ancient Greek, *akr-* is the root of the noun *akra* (or *akron*) ‘top, extremity, edge’, and of the adjective *akr-os/-a/-on* ‘high, extreme’.⁸ Like other lexemes, *akr-* participates in compounds (in this case, [N N] or [A N] ones), as in the following examples, where a linking element/compound marker *-o-* appears between the root and the second constituent:⁹

(1) Ancient Greek

a. *akr-o-xlieros* < *akr(a/on) xlieros* (Hippocrates, 5th c. BC)

‘little warm’ edge warm

b. *akr-o-polis* < *akr(a) polis*

‘high town’ high town

According to Babiniotis (1969: 111), the formations with *akr-* have been subject to a semantic drift, the first indications of which go up to the 8th c. BC (2a), where *akr-* seems to quantify the meaning of the base by bringing either a weakening (2a) or an intensification (2b).

(2) a. *akr-o-knephaios* (Hesiodus, 8th c. BC) < *akr-* kneph(as)

little dark edge cloudness/darkness

b. *akr-o-mane:s* (Herodotus, 5th c. BC) < *akr-* -mane:s¹⁰

very mad extremity mad

In Hellenistic Koine (ca 3th c. BC – 3th c. AD), the examples of this use become more frequent, where *akr-* appears mainly in contexts, where the meaning of the base is weakened. However, compounds with *akr-* bearing the original meaning of *akr(a/on)* are still common:

(3) a. *akr-o-karpos* (Theophraste, 4th c. BC) < *akr(on) karp(os)*

with fruits at the top top fruit

b. *akr-o-lith(os)* (Palatine anthology, 5th c. AD) < *akr(on) lith(os)*

with stone edges edge stone

lexical items are categorically underspecified, and get their categorial specifications by appearing in syntactic structures.

⁷ See Hopper (1991) for the gradable nature of grammaticalization in general.

⁸ Most adjectives in both Ancient and Modern Greek have three forms (corresponding to distinct inflectional paradigms), depending on their gender value, i.e. masculine, feminine, and neuter.

⁹ See Ralli (2008b) for the notion of compound marking and compound markers in Greek.

¹⁰ *-man(es)* is a bound nominal form, deriving from the verb *main(omai)* ‘to be in a rage’.

Examples demonstrating the weakening function of *akr-* are multiplied during the late medieval period (around the 12th c. AD), where there are also instances of verbal formations:

- | | | |
|---|---------|-------------|
| (4) a. akroeksispazo (Glykas, 12 th c. AD) | < akr- | eksispaz(o) |
| to shake a bit | | to shake |
| b. akrioxtipo (Chronicle of Moreas, 14 th c. AD) | < akri- | xtipo |
| to softly knock | | to knock |
| c. akralafrono (Pseudo Georgil, 15 th c. AD) | < akr- | alafrono |
| to lighten a bit | | to lighten |

Today, it still appears in certain dialectal areas, frequently in Crete, and sporadically in Cyprus, Pontus, South Italy, the Dodecanesian islands, Thrace, and the Peloponnese.

- | | | |
|---|----------|--------------|
| (5) a. akrokuzulizo (Crete) | < akr- | kuzulizo |
| to softly distract | | to distract |
| b. krofoume (Cyprus) | < (a)kr- | fou(me) |
| to be a bit afraid | | to be afraid |
| c. akriokitrininos (Peloponnese) | < akri- | kitrininos |
| yellowish | | yellow |
| d. akranixtos (Pontus, South Italy, Dodecanesian islands, Thrace) | < akr- | anixtos |
| a bit open | | open |

Crucially, while the ancient noun root *akr-* is attached to nominals (adjectives and nouns), the dialects display many verbal examples with *akr-* as first constituent. This is an indication that it has become neutral with respect to the category of the base it combines with, and thus, argues in favor of a possible prefixal status. In fact, as pointed out by Amiot (2005: 184), the ability to combine with different categories of lexemes can be a criterion according to which we may distinguish a lexical item from a prefix.

It is also important to add that the prefixal status entails a form restructuring, from *akr-* to [*akr- + -o-*], since the linking element *-o-* would no longer be considered as a compound marker, and it should be analyzed as being incorporated onto the prefix. Note that this collapsing together of adjacent forms has been proposed by Lehmann ([1982] 1995) to be one of the parameters for grammaticalization (coalescence). However, beside the merger of the root and the linking element, there is no other substantial form change. For instance, in Cretan, where *akro-* is very productive (6a), it keeps its original form. A slight change is observed in Cypriot and Peloponnesian, where *akro-* appears as *kro-* (6b) or *akrio-* (6c), respectively:

- | | | |
|---------------------------------|-----------|---------|
| (6)a. akrovoitho (Cretan) | < akro- | voitho |
| to help a bit | | to help |
| b. krolalo (Cypriot) | < (a)kro- | lalo |
| to have a small talk | | to talk |
| c. akriokokinos (Peloponnesian) | < akrio- | kokinos |
| little red | | red |

It should be noticed that the change in Cypriot (*kro-* in 6b) is triggered by the application of a general phonological law applying to certain dialects, according to which unstressed vowels are usually deleted at the beginning of words (cf. Newton 1972). Note that the Peloponnesian *akrio-* (6c) is a particularly interesting case, since it establishes a formal link with the Medieval word types *akri* and *akria* 'edge'¹¹, which coexist with the Classical Greek form *akra*. *akrio-* is firstly detected in the 14th century, as illustrated by the examples of the Chronicle of Moreas in (4b), and can be used as an indication that the prefixization of *akr(a/i/ia)* into *akro-* (or *akrio-*, depending on the area) has occurred by that period. Since *akrio-* comes from *akri/akria*, it does not involve any real form change.

Crucially, parallel to the use of *akro-* as a meaning quantifier, the noun forms *akri/akria* 'edge' or *akro* have never disappeared from the language, as shown by the following

¹¹ *akra* was the Attic form, while *akri* was the Ionian one.

Standard Modern Greek examples, where they keep their original meaning. They still form compounds (7a) or show as free items in syntactic structures (7b):

- (7)a. akrokeramo < akr(o) keram(idi)
 tile of the edge edge tile
 b. I akri/to akro/i akria tu dromu
 The edge of.the road

Finally, it is important to point out that the coexistence of the old noun with the new prefix does not pose any problems for the prefixation hypothesis: it illustrates a typical case of ‘divergence’, which is justified and accounted for within the framework of grammaticalization theory (Hopper 1991: 11)¹².

2.2 *moro-*

mor- is the root of the Ancient Greek adjective *mo:r-os/-a/-on* ‘idiot, silly’, and with this meaning appeared in a small number of nominal compounds of the classical period:

- (8) mo:r-o-logos (Aristotle, 4th c. BC) < mo:r- -log(os)¹³
 who talks silly silly who talks

As Babiniotis (1969: 154) notes, in the 12th century, formations with *mor-* display traces of a hypocoristic function. For instance, *moroipnos* in (9) is ambiguous: it may mean a ‘silly sleep’, where *mor-* keeps the original meaning, but also ‘little sleep’:

- (9) moroipnos < mor- ipnos (Glykas, 12th c. AD, 170 TLG)
 little sleep sleep

However, there is no other evidence of this hypocoristic meaning in the subsequent centuries, and it is only in the 17th c. AD, where the first examples of a similar use are detected in a chronograph from Serres, a town in the northern part of Greece (Macedonia):

- (10)a. morogematos < mor- gematos
 not very full full
 b. moropsaltis < mor- psaltis
 who knows some chanting chanter
 c. morofovume < mor- fovume
 to be a bit afraid to be afraid

What is crucial about these occurrences is not only the new hypocoristic meaning, which at least for the examples (10a,b) is not transparent to the original meaning (‘silly’), but also the fact that *mor-* can be added to verbs (10c). This property to combine with lexemes of various categories leads us to suppose that combinations with *mor-* are not compounds, but derived words, i.e. prefixed words. Were *mor-* an adjective, the only possible combinations would have been those with a nominal base, such as the ones that we find in earlier texts. We further suppose that the prefixal use also leads to a form restructuring (coalescence), from *mor-* to *moro-*, as we have supposed for *akro-*, according to which *-o-* is no longer a compound marker, but a prefix final vowel.

Today, the prefix *moro-* can be found in dialects all over Greece (11), but the number of occurrences is very restricted, and the native speakers of these dialects cannot create productively new formations:

- (11)a. moroskotina (Mykonos) < moro- skotina
 little dark dark
 b. morovrasto (Kythera) < moro- vrasto
 little boiled boiled
 c. moranixtos (Chios) < moro- anixtos
 bit open open

¹² “When a lexical form undergoes grammaticization to a clitic or affix, the original form may remain as an autonomous element, and undergo the same changes as ordinary lexical items.”

¹³ *-log(os)* is a bound nominal, which derives from the verb *lego* ‘to talk’. See Ralli (2008a) for more information about those nominal elements.

- d. moroprasinizo (Macedonia) < moro- prasinizo
to become little green to become green
- e. morokegome (Epeiros) < moro- kegome
to be a bit burnt to be burnt
- f. morovrexī (Euboea) < moro- vrexī
to rain a bit to rain

The limited number of these occurrences and the lack of productivity of prefixing *moro-* to other lexemes, drive us to the conclusion that *moro-* came close to become a prefix in some parts of Greece. However, for some reason it disappeared, leaving certain examples, such of those in (11) as fossilized cases.

It is also important to add that with respect to its form, *moro-* has not undergone any specific phonological changes across centuries, with the exception of the shortening of the ancient root vowel /o:/, which, however, has affected all Greek long vowels in the Hellenistic period (ca. 3rd c. BC - 3rd c. AD).

At this point, it is worth noticing that parallel to the appearance of the hypocoristic function in Serres (17th c. AD), in two other areas, Cyprus and Crete, the adjective *mor(os)* 'silly' seems to have undergone a recategorialization as noun, with the meaning of 'baby'. This noun is found as a free item in syntax (12a,b), and as a stem constituent of [N N] Cretan compounds (12c) in various texts dating of the 16th and 17th centuries:

(12)a. Cretan (Erotokritos A2239, 17th c. AD)

San to moro opu kianis fajto ðen t' arminevji ke kin ot ora jeniθi na vri vizi jirevji

Lit. Like the baby that nobody food NEG it recommend.3Sg and it any time bear.3PassSg PRT look.3Sg for breast

'As for the baby for whom nobody recommends any food, but by the time he is born he looks for breast-feeding'

b. Cypriot :

moron pedin (Poèmes d'amour, 16th c. AD)

baby child

c. Cretan

morokopelo (Stathis, 17th c. AD)

young man

The noun *moro* spread all over the Greek speaking world, since it is part of today's common vocabulary, while its ancestor *mor(os)* 'silly' has disappeared from the common language.¹⁴ However, compounds with the stem of the noun *moro* as one of their constituents are not rare, especially in the dialects of Lesbos and Aivali¹⁵, as the following examples illustrate:

(13) Lesbian / Aivaliot

a. mur-u-klegu¹⁶ < mur- klegu

to cry like a baby baby to cry

b. mur-o-panu < mur- pan(i)

baby cloth baby cloth

Interestingly, in the dialect of Apiranthos of the island of Naxos (14), which is related to Cretan, the noun *moro* seems to have developed a new evaluative function. In this dialect,

¹⁴ It subsists only in some expressions of a very formal type of language (in the so-called 'katharevousa'), which are reminiscent of Ancient Greek.

¹⁵ Aivaliot is the Asia Minor dialect of the former Greek-speaking town of Kydonies (also called Aivali), today's Ayvalik, till 1922. This dialect is still spoken in certain dialectal enclaves in Greece, which are inhabited by first, second, and third generation refugees, who have settled there after the end of the war between Greece and Turkey, and the Lausanne treaty in 1923.

¹⁶ -u- is the linking vowel/compound marker. It is an underlying /o/ which has become /u/ in unstressed position because of a dialectal phonological law applying to the northern Greek dialects, among which those of Lesbos and Aivali, which raises the mid unstressed vowels /o/ and /e/ into /u/ and /i/ respectively.

there are formations, where the stem of *moro* can be seen as a diminutivizer of the meaning of the base:

(14) Apiranthos (Naxos)

- a. moragatho < mor- agath(i)
 little thorn little thorn
 b. moromagazo < mor- magaz(i)
 little shop little shop

However, this evaluative use should be considered as an independent development from that of the Serres dialect. Beside the fact that the new diminutivizing *mor-* appears in a different area from that of Northern Greece (Naxos and Crete are located in South Greece), it is attached only to nouns, and its semantic and formal relation with the new noun formation *moro* 'baby' is very transparent. Therefore, it is legitimate to assume that the Apiranthos *mor-* still retains its lexical character, and has not acquired the prefixal status.

2.3 *sjo/so-*

sjo-/so- (< *sio-*) as first constituent of morphologically complex words originates from the adverb *isja* (< *isia*) 'straight'. It appears under the form of *sjo-* in Western Crete, while in the eastern part of the island, an independently motivated palatalization law reduces *sjo-* into *so-*. In the early texts of the 16th and 17th centuries, the original adverbial stem *is(i)-* is a compound constituent, as illustrated by the examples in (15). In these examples, the unstressed initial vowel /i/ is deleted, due to a phonological law erasing initial unstressed vowels, as already mentioned for *akro-*, and a compound marker *-o-* appears between the two compound constituents:

(15)a. Ta kanu ki apomenusi me texni s-o-themena (Panoria A 416)

- Lit. Them make.1SG and remain.3PL with art straight-put
 'I make them and they remain as such with an artistic straight manner'
 b. s-o-pato horafi (Varuchas, notary. 1598.353.2)
 Lit. straight-stepped land
 'flat land'

Dimela (2005) and Ralli & Dimela (to appear) have shown that parallel to the original word where it came from, *sjo-* is used in today's Cretan as an intensifying prefix, and is attached to several categories, i.e. to verbs (16a), adjectives (16b), adverbs (16c), and nouns (16d):

(16) Cretan

- a. sojerno < so- jerno
 to become very old to become old
 b. soaspros < so- aspros
 very white white
 c. sodreta < so- dreta
 very straight straight
 d. sogopanisma < so- kopanisma
 thrash walloping

sjo- is very frequent, and participates in the creation of everyday neologisms, some of which cannot be found in the most updated Cretan dictionaries (e.g. Idomeneas 2006 and Ksanthinakis 2000). For instance, Dimela (2005) reports the verb *sjoksejivedizo* 'highly humiliate', which has been produced by native speakers during her field work.

The prefixal status of *s(j)o-* is also proved by the fact that, on synchronic grounds, native speakers make no link between its initial lexical meaning of 'straight' and the actual intensifying function. For instance, they often mix up *s(j)o-*, originating from *isja* 'straight', with a prefix *sin-* (from the Ancient Greek preposition *sin* 'together, plus' cf. Charalabakis 2001). Following Dimela (2005), this confusion is due not only because *sjo-* and *sin-* are not very distant phonologically, but also because among the interpretations of their

morphologically produced words there is a notion of similarity. The first traces of such a mixing can be detected as far as to the 17th c. AD. Consider the following examples:

(17) a. k'i djo so-bropatusasi (Erotokritos A 37)¹⁷

Lit. and.the.two straight-stepped.3PL

'And both of them have the same age'

b. sjotseros < sjo- ker(os) / sigeritis < sin ker-itis
of the same age weather/time / of the same age time/weather-DAFF

(17a) is ambiguous with respect to which of the two, (*i*)*sjo*- or *sin*- is used: formally, the first constituent *so*- appeals to the original *isja*. However, the fact that the initial consonant of the base (*propato* or *porpato* 'to walk') becomes a voiced /b/ shows that the previous constituent ends in a nasal /n/, which belongs to *sin*-. The mixing is further demonstrated by (17b), where without any change in the meaning, the same base is prefixed by either *sjo*- or *sin*-. Further proof is found in the files of the *Centre of Research of Modern Greek Dialects* of the *Academy of Athens*, where the verb *sofiliazio* (< *filiazo*¹⁸ 'apply') is given two different interpretations: in certain files, *so*- is attributed to the word *isja*, while in others, an anonymous lexicographer claims that it comes from the preposition *sin*.¹⁹

Crucially, as noted by Ralli & Dimela (to appear) and Ralli (2009b, 2010), in some northern dialects, mainly in Lesbian and Aivaliot, a corresponding item *sa*-, also originating from the adverb *isja*, appears preposed to locative adverbs.

Consider the examples in (18):

(18) *sapera* 'far away' < *sa*- *pera* 'away'
sadju 'over here' < *sa*- *edju* 'here'
saki 'over there' < *sa*- *iki* 'there'
sakatu 'straight down there' < *sa*- *katu* 'down'
sapanu 'straight up there' < *sa*- *apanu* 'above'
samesa 'more inside' < *sa*- *mesa* 'inside'

Ralli & Dimela (to appear) have shown that, contrary to Cretan speakers, all native speakers of Lesbian and Aivaliot are aware of the relationship that *sa*- bears with the original word *isja*. In these dialects, the fact that *sa*- is still semantically transparent with respect to *isja* casts doubt on the hypothesis that *sa*- is a real prefix. If it is a lexeme, its combination with the locative adverbs could be analyzed as an instance of compounding. In fact, *sa*-, under its full adverbial form *isa*, also appears at the right-hand position of adverbial compounds, as for instance, in the following formation:

(19) *uloisa* 'all straight' < *ulu* 'all' *isa* 'straight'²⁰.

Moreover, the appearance of *sa*- in morphologically complex adverbs is of limited productivity, since it is restricted to a handful of examples containing specific locative adverbs, as illustrated by the ungrammatical example of **saksu* in (20):

(20) **saksu* 'more outside' < *sa*- *oksu* 'outside'

Finally, like *sjo*-, *sa*- has undergone a phonological change with an initial /i/ deletion and the internal loss of the semi-vowel /j/ (palatalization). However, both phonological changes are due to general phonological laws, which apply to several Modern Greek dialects, independently of the particular morphological environment of the *s(j)o*-/*sa*- formations.

¹⁷ Literary texts of the 17th century are written in the dialectal variant of Eastern Crete, where the prevalent form is *so*-. Therefore, it is not a coincidence that *so*- is phonologically confused with *sin*-, since it is more similar with the latter than its variant *sjo*-.

¹⁸ The verb either comes from *thiliazio* (< *thilia* 'noose, eyelet') or is of an unknown etymology.

¹⁹ Interestingly, a number of comparable cases can be shown in the dialects of Cyprus, a number of Cycladic islands (e.g. Naxos, Thera), Euboea and Samos, although not with the same frequency.

²⁰ In this case, there is no need for /i/ deletion, since /i/ is not in the initial position.

3 Discussion

As seen with the data above, there is no doubt that *akro-* in several dialects, and *s(j)o-* in Cretan are prefixes resulting from a prefixization process. The evolution of *akro-* cuts across the history of Greek, since the first indications of a semantic change appeared in the early years of the historical period (8th c. BC), while *s(j)o-* is a recent formation. *moro-* is a different case: there are traces of a prefixal use in the post medieval period (17th c. AD), but, as shown in 2.2., this use has disappeared from the language. In contrast, the original adjectival lexeme (with the meaning of ‘silly’) has been nominalized (with the meaning of ‘baby’), and from that, a new evaluative use seems to be under development, especially in the dialect of Apiranthos. However, this new form is not a true prefix yet, since the connection with its source is quite transparent on both semantic and structural grounds. For instance, it is significant that *moro-* as a diminutivizer cannot combine with adjectival bases, as opposed to other diminutive affixes in Greek, which can be attached to both nouns and adjectives.

The status of a real prefix is doubtful with respect to the Lesbian/Aivaliot *sa-* too, which is also transparently linked to its source, and has specific combinatorial properties, since it is combined with a small number of locative adverbs.

As mentioned in section 1, prefixization is an instance of morphologization, and its realization is due to a number of parameters. With respect to the two general parameters that are usually assumed to be involved not only in morphologization, but in every grammaticalization process, that is the phonological and the semantic ones, our data have shown the following two facts:

a) Phonological erosion may precede or follow resemanticization, but it is not a compulsory criterion for an item to be morphologized. In fact, we have seen that *akro-* has become a prefix in a number of Modern Greek dialects, without being subject to any phonological change, and that the slight change that is attested with respect to the Cypriot *kro-* is not related to the process of prefixization itself but is due to a general phonological law. Nevertheless, the Cretan *s(j)o-* proves that phonological change, although independently motivated, is part of the prefixization process of the adverb *isja*, as it led to the confusion with the preposition *sin* (see section 2.3).

b) Resemanticization has affected all three examined items. Our data have provided support to Booij’s (2005) statement that semantic change precedes the formal one. Nevertheless, as shown by *moro* in the dialect of Apiranthos, and by *sa-* in Lesbian and Aivaliot, resemanticization is not a sufficient parameter to ensure completion of prefixization. In fact, those two items are still close to lexemes, and speakers still link them to their sources.

In section 1.2, we have put forward the hypothesis that since compounding and prefixation are morphological processes, at least for Greek, the decisive criteria for an item to become a prefix should be morphological. With few exceptions, researchers agree that one of these criteria refers to the property of boundness (see, among others, Booij 2005). However, as already mentioned in section 1.1., in a language like Greek, both prefixes and the first constituents of compounds are bound, the latter being stems deprived of their inflectional endings. Seen from this perspective, the distinction between the first constituent of a compound and the prefix of a prefixed word should not be based on the non-separability property of these items. In fact, van Goethem (2007) has also reached the same conclusion in her examination of Dutch preverbs. In this paper, we would like to suggest that the application, or non-application, of this parameter should be viewed as being language dependent. In Greek, prefixes have a ‘more bound character’ than stems, since they do not appear in syntactic constructions as free items (stems can be used as free words with the appropriate inflectional endings).²¹ In this sense, the non-separability

²¹ Some Modern Greek prefixes though share the same form with prepositions that are free items. For instance, the prefix *apo*, in a prefixed verb like *apografo* ‘to record’, has the same form with the

criterion is not irrelevant to a prefixization process in this language, since it makes a morphologized item to gain a greater degree of boundness. Nevertheless, we would also like to claim that it should be considered as a criterion for distinguishing a Greek prefix from a non-prefix, and not as a parameter, which may be directly involved in a prefixization process.

The same considerations apply to the decategorialization of an item, and the property of occupying a specific position within a morphologically complex word. A constituent which is category neutral, and appears at the left-hand position²² of a word, has already become prefix, as opposed to stems, which belong to specific grammatical categories, and may appear as first or second items, depending on the case. Therefore for a particular item, boundness, decategorialization, and fixed position are strong indicators of a prefixal status. These properties signal the final stage (the result) of prefixization, and should not be viewed as parameters, which may induce the item to become a prefix.

The question that still requires an answer though concerns the parameters which are typical of a prefixization process, and characterize the incipient stages, where variable phenomena occur. At this point, we would like to propose that the decisive factors for the completion of a prefixization process are a) the expansion of the combinatorial properties of an item (in accordance with Amiot 2005), and b) the raise of productivity of a candidate prefixation pattern. For instance, in Ancient Greek, *akr-* and *mor-* are attached to nouns to form compounds. In contrast, much later (*akro-* around the 12th century and *moro-* at the 17th century) the two items appear to be combined with nouns, adjectives and verbs. In other words, they have become category neutral, like true prefixes. However, while formations with *akro-* have been multiplied, and since the 12th c. are massively used in a number of dialects, those with *moro-* have disappeared. The spread of the *akro-* formations, and the disappearance of those with *moro-*, are mainly due to the degree of productivity according to which their combining processes occur. In fact, as shown in section 2.2, occurrences with *moro-* are found only in a single 17th century document from Serres. Low productivity prohibits the use of *moro-* to spread, and thus, its prefixal status is doubtful.

The same considerations apply to *s(j)o-*: we have seen in 2.3 that *s(j)o-* after being confused with the prefix *sin-* (around the 17th century) there is a significant raise of productivity of the process. We suggest that category neutrality, as well as the high productivity of attaching *s(j)o-* to several bases has induced it to emerge as a real prefix.

Nevertheless, as also seen in section 2.3, there is no sufficient justification for the hypothesis that its cognate Lesbian and Aivaliot *sa-* is a prefix. Given the unclear status of *sa-*, we may suppose that it is in the process of losing its lexeme independence, and thus, it may be considered as a kind of prefixoid. Although there are certain indications (e.g. form reduction and extended meaning), which suggest a morphologization in progress, there is no guarantee that it will result into being one: for instance, it shows no expansion of its combinatorial properties, being combined only with certain locative adverbs. It is important to point out that *sa-* illustrates the intermediate stage of a prefixization cline, where true prefixes occupy one pole, lexemes the other pole, and prefixoids are situated in between (cf. Bauer 2005, and Ralli 2010 for the notion of cline). Thus, it confirms the general claim that grammaticalization changes are accomplished gradually, as proposed by many linguists (see, among others, Meillet 1912, Lehmann 1985, Lichtenberk 1991).

preposition *apo* denoting the origin (e.g. *Erxome apo tin Athina* 'I come from Athens'). In accordance with Ralli (2005), we consider the prefix *apo* to be a bound item, and distinct from the preposition.

²² Note, however, that detecting the exact position of constituents in morphologically complex words requires an accurate documentation, something which is very difficult to have if one deals with diachronic sources, where crucial evidence is often missing (c.f. Manolessou 2008).

4 Conclusions

In this paper, we have shown that the properties of *akro-*, *moro-* and *sjo/so-* provide significant insights about the nature of a prefixation process, which implies a greater morphological involvement, since it belongs to morphologization. With respect to the general parameters playing a role in a grammaticalization process, i.e. phonological erosion and resemanticization, we have argued that they cannot ensure completion, as the crucial parameters of prefixation have to be morphological. We have proposed that decategorialization, boundness, and positioning signal the final stage (result) of prefixation and are not directly linked to the process itself, since the specific morphological parameters leading to completion are: a) the expansion of the combinatorial properties, and b) the raise of productivity of a word-formation process.

Finally, elaborating on dialectal data, we have claimed that dialects provide crucial evidence for our argumentation, evidence usually absent from the standard form of a language.

References

- Amiot D. (2005). Between Compounding and Derivation: Elements of Word-formation Corresponding to Prepositions. In W. Dressler, D. Kastovsky, O. Pfeiffer and F. Rainer (eds.), *Morphology and Its Demarcations*. Amsterdam: John Benjamins. 183-195.
- Anderson S. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.
- Babinotis G. (1969). *O Dia Sintheseos Ipokorismos is tin Ellinikin* [Hypocoristic Formation via Compounding in Greek]. Athens.
- Bauer L. (2005). The Borderline between Derivation and Compounding. In W. Dressler, D. Kastovsky, O. Pfeiffer and F. Rainer (eds.), *Morphology and Its Demarcations*. Amsterdam: John Benjamins. 97-108.
- Booij G. (2005). Compounding and Derivation: Evidence for Construction Morphology". In W. Dressler, D. Kastovsky, O. Pfeiffer and F. Rainer (eds.), *Morphology and Its Demarcations*. Amsterdam: John Benjamins. 109-132.
- Charalambakis C. (2001). Prolegomena se ena Etymologiko Lexico tis Kritikis Dialektou [Preface of an Etymological Lexicon of the Cretan Dialect]. *Kritologika Meletimata : Glossa, Logotechnia, Politismos*. Irakleio: Panepistimiakes Ekdoseis Kritis. 97-109.
- Dimela E. (2005). *I prothimatopoiisi stin Kritiki Dialekto. I Periptosi ton sin-, s(j)o-, (e)so-* [Prefixation in the Cretan Dialect. The Case of *sin-*, *s(j)o-*, *(e)so-*]. University of Athens: Master Thesis.
- Fleischer W. (1969). *Wortbildung der Deutschen Gegenwartssprache*. Leipzig: VEB Bibliographisches Institut.
- Gaeta L. (1998). Some remarks on analogy and grammaticalization. In A. Giacalone-Ramat & P.J. Hopper (eds.), *The Limits of Grammaticalization*. Amsterdam/ Philadelphia: John Benjamins Publishing Company. 89-105.
- Givón T. (1971). Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistic Society* 7(1): 394-415.
- Givón T. (1979). *On Understanding Grammar*. New York: Academic Press.
- Heine B. (2003). Grammaticalization. In B.D. Joseph & R. D. Janda (eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell. 575—601.
- Heine B. & T. Kuteva (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine B. & T. Kuteva (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine B. & T. Kuteva (2007). *The Genesis of Grammar: a reconstruction*. Oxford: Oxford University Press.
- Heine B. & M. Reh. (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hopper P.J. (1991). "On Some Principles of Grammaticalization". In E.C. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*. Amsterdam / Philadelphia: John Benjamins Publishing Company. 17-35.
- Hopper P.J. & E.C. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Iacobini C. (2004). Prefissazione. In M. Grossman & F. Rainer (eds.), *La Formazione delle Parole in Italiano*. Tübingen: Max Niemeyer Verlag. 97-161.

- Idomeneas M.I. (2006). *Kritiko Glossary* [Cretan Glossary]. Irakleio: Vikelaia Vivliothiki.
- Joseph B.D. (2003). Morphologization from Syntax. In B.D. Joseph & R.D. Janda (eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell. 472-492.
- Klausenburger J. (2002). Grammaticalization within a theory of morphocentricity. In I. Wischer & G. Diewald (eds.), *New Reflections on Grammaticalization*. Amsterdam/ Philadelphia: John Benjamins Publishing Company. 31-43.
- Ksanthinakis A.B. (2001). *Leksiko Ermineftiko kai Etumologiko tou Ditikokritikou Glossikou Idiomatos: sunonuma – antitheta, epeksigimatika paradigmata idiomatikou logou* [Interpretative and Etymological Lexicon of Western Cretan dialect]. C. Charalabakis (ed.). Irakleio: Panepistimiakes Ekdoseis Kritis.
- Lehmann C. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e stile* 20 (3): 303-318.
- Lehmann C. ([1982] 1995). *Thoughts on Grammaticalization*. München – Newcastle. Lincom Studies in Theoretical Linguistics 01.
- Lichtenberk F. (1991). On the gradualness of Grammaticalization. In E. C. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*. Amsterdam/ Philadelphia: John Benjamins Publishing Company. 1: 37-80.
- McMahon A. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Manolessou I. (2008). On historical linguistics, linguistic variation and Medieval Greek. *Byzantine and Modern Greek Studies* 32 (1): 63-79.
- Marchello-Nizia C. (2006). *Grammaticalisation et changement linguistique*. Bruxelles: De Boeck.
- Meillet A. (1912). L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)* 12(26): 6. Repr.: Meillet 1921: 130-148.
- Newton B. (1972). *The Generative Interpretation of a Dialect. A study of Modern Greek Phonology*. Cambridge: Cambridge University Press.
- Ralli A. (2005). *Morfologia* [Morphology]. Athens: Patakis.
- Ralli A. (2007). *I Sinthesi Lekseon: Diaglosiki Morphologiki Prosengisi* (Compounding: a Cross-linguistic Morphological Approach). Athens: Patakis.
- Ralli A. (2008a). Greek Deverbal Compounds with Bound Stems. *Southern Journal of Linguistics* 29 (1-2): 150-173.
- Ralli A. (2008b). Compound Markers and Parametric Variation. *Language Typology and Universals (STUF)* 61: 19-38.
- Ralli A. (2009a). Hellenic Compounding. In R. Lieber & P. Stekauer (eds.), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 453-463.
- Ralli A. (2009b). Morphology meets Dialectology. *Morphology* 19 (1): 87-105.
- Ralli A. (2010). Compounding versus Derivation. In S. Scalise and I. Vogel (eds.), *Cross-Disciplinary Issues in Compounding*. Amsterdam: John Benjamins. 57-74.
- Ralli A. (in preparation). *Greek Compounds*. To appear at Springer.
- Ralli A. & E. Dimela (to appear). On the Borderline between Derivation and Compounding: The *sa* adverbs in the Dialectal Varieties of Lesbos, Kydonies (Aivali) and Moschonisia. In P. Pavlou et al (eds.), *Proceedings of the 3rd International Conference on Modern Greek Dialects and Linguistic Theory*. Nicosia, Cyprus.
- Ramat P. (2001). Degrammaticalization or transcategorization?. In C. Schaner-Wolles, J. Rennison and F. Neubarth (eds.), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Torino: Rosenbach and Sellier. 393-401.
- Stekauer P. (2005). Compounding and Affixation. Any Difference?. In W. Dressler, D. Kastovsky, O. Pfeiffer and F. Rainer (eds.), *Morphology and Its Demarcations*. Amsterdam: John Benjamins. 151-159.
- ten Hacken P. (2000). Derivation and Compounding. In G. Booij, C. Lehmann & J. Mugdan (eds.), *Morphologie - Morphology: Ein Handbuch zur Flexion und Wortbildung - A Handbook on Inflection and WordFormation*. Berlin: Walter de Gruyter. 1: 349-360.
- van Goethem K. (2007). The Boundary Between Composed and Derived Verbs Revisited: Dutch Preverbs on a Cline of Grammaticalization. *MorphOn/Virtual Yearbook 2007*: www.morphologyonline.com
- van Goethem K. (2008). "Oud-leerling versus ancien élève: A Comparative Study of Adjectives Grammaticalizing into Prefixes in Dutch and French". *Morphology* 18: 27-49.
- van Marle J. (1985). *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.

Loan verb adaptation in Greek dialectal variation: A first approach¹

DIMITRA MELISSAROPOULOU

University of Patras

dmelissa@upatras.gr

Abstract

The paper investigates loan verb borrowing and adaptation in the light of the evidence provided by Greek dialectal variation. Examining the mechanisms and paths via which verbs can be borrowed and adapted in Greek dialectal systems, according to Wichmann & Wohlgemuth (2008) typological classification of loan verb accommodation strategies, we argue in favour of the prevailing influence of structural factors (i.e. productivity of the selected pattern, base specifications and phonological equivalences) to the selection of -a- specific accommodation mechanism-s- across dialectal varieties.

1. Introduction

Lexical borrowing as well as adaptation of loans is a favorite topic among linguistic studies, both for theoretical and applied reasons i.e. understanding the nature of language change via the identification of the constraints language is subject to, and using the constraints for the reconstruction of unattested language change and language situations (cf. Haspelmath 2008).

Several claims regarding borrowability have been made the most important of which, for the purposes of the present paper are the following: a) lexical items are more likely to be borrowed than grammatical items and words are more likely to be borrowed than bound morphemes (cf. Moravcsik 1978, Field 2002) and b) different spheres of the vocabulary are borrowed more easily, while others significantly less easily.

According to Hock & Joseph (1996:257) basic vocabulary, referring to essential human activities, e.g., *eat, sleep, do, have, be* is the more resistant sphere. Moreover, it is a general assumption that nouns are borrowed more easily and thus preferentially than other parts of speech (see among others Whitney 1881, Moravcsik 1978, Myers - Scotton 2002), since according to Myers - Scotton (2002: 240), “[...] *they receive, not assign, thematic roles*”, *so their insertion in another language is less disruptive of predicate – argument structure*”².

In terms of contact, Dawkins(1916:197), focusing on Asia Minor Greek, had already claimed that “[...] *verbs are borrowed much less easily than other parts of speech and only appear in any number when the vocabularies of two languages have reached a high degree of fusion...[...] often to the complete exclusion of their Greek equivalents.*” A more strong thesis is that of Moravcsik (1978: 111) who argues that a “[...] *lexical item whose meaning is verbal can never be included in the set of borrowed properties*”³.

The aim of this paper is to investigate loan verb borrowing and adaptation⁴ in the light of the evidence provided by Greek dialectal variation (i.e. Pontic, Cappadocian,

¹ The author wishes to thank the Greek State Scholarships Foundation for funding the present work.

² The same claim is made by Van Hout & Muysken (1994) based on the Quechua language.

³ Additionally, “[...] *if verbs are borrowed, they seem to be borrowed as if they were nouns: the borrowing language employs its own means of denominal verbalization to turn the borrowed forms into verbs before using them as such*” Moravcsik (1975, 1978: 111-112).

⁴ The variety of terms, i.e. loanword adaptation, accommodation, integration, assimilation used in the literature are considered to be synonymous and thus are interchangeable in this study,

can be a verb root or an inflected type. The authors give an example of verb adaptation in Figuiq Berber from French, which can be seen in the example below:

- | | |
|---------------------------------------|------------|
| Figuiq Berber | < French |
| (3) i-gōfla [3SG.- be swollen. PERF.] | < gonfler |
| ‘he is swollen up’ | ‘to swell’ |
- (Data from Wichmann & Wohlgemuth 2008)

4) The last accommodation strategy is the inflectional transfer. In this case, the loan verb is not accommodated in the morphology of the target language. On the contrary, it carries its verbal morphology from the language source maintaining its functions in the new system. Example of inflectional transfer is found in Agia Varvara variety of Romani where the borrowed verb type *okursun*, shown under (4), carries its Turkish inflectional marker *-sun*:

- | | |
|---------------------------|------------------------|
| (4) Romani (Agia Varvara) | < Turkish ⁷ |
| okursun | < okurmak |
| [read.2SG.] | < okurmak ‘to read’ |
- (Data from Bakker 2005:9 in Wichmann & Wohlgemuth 2008)

Wichmann & Wohlgemuth acknowledge the fact that the borrowing pattern which the language target will adopt is related to its structural characteristics. However, they claim that the existence of more than one mechanism proves that the ‘structural outcome’ cannot be predicted on structural terms. Moreover, they form the hypothesis that the existence of different accommodation patterns in the target language correlates to the degree of exposition to the source language. In this spirit, they propose these strategies to form a hierarchy to be tested, according to which the lowest accommodation grade is related to the light verb strategy, a some how higher grade is marked by indirect insertion while, in the case of direct insertion there is no accommodation effort, acknowledging a special status to it. Finally, they suggest that the relative change in the accommodation strategy used by the target language is related to the relative degree of bilingualism in the source language.

3. Accommodation strategies in the dialects in study

As stated in the introduction, we focus on the adaptation of verb loans in different dialectal systems from the same language source. The language source is Turkish, which is an agglutinative language of the Altaic family and the three dialects in study, Cappadocian, Pontic and Aivaliot, varieties of Greek, which is a fusional language and member of the Indo-European family. Pontic, Cappadocian and Aivaliot were once spoken in the Ottoman Empire (in the areas of Northwest Turkey, Cappadocia, and West Turkey, respectively). After the end of the war between Greece and Turkey in 1922, the dialects continue to be spoken in Greece, within communities of first, second and third generation refugees⁸. Let us see the accommodation mechanisms in use in each dialectal system.

3.1 Cappadocian

⁷ According to Bakker (2005) Turkish loan verbs are inflected with the Turkish suffixes in present and past tense except for the 1.Pl. suffix of the past.

⁸ Pontic is it is still spoken by an unknown number of Pontic Muslims who still live in the same area in Turkey (see Mackridge 1999, Drettas 1999, 2000, Kaltsa and Sitaridou this volume, Michelioudakis & Sitaridou this volume).

Cappadocian⁹ is often used in the literature as a prototypical example of ‘heavy borrowing’ in terms of Thomason & Kaufman’s borrowing scale, referring to ‘overwhelming long-term cultural pressure’ (Thomason & Kaufman 1988:50). The length and intensity of cultural and linguistic contact led Dawkins to the following statement about Cappadocian dialect “[...*the body ha[d] remained Greek but the soul ha[d] become Turkish...*”], Dawkins (1916:198). It should be noted that although Cappadocian is a variety of Greek origin and its basic morphological structure is fusional, it displays hints of agglutinative patterns due to language contact with Turkish. More importantly it is the only variety where agglutinative inflectional structures are attested (cf. Dawkins, 1916 and Janse, forthcoming). Lastly, the Cappadocian dialect is subdivided into two basic groups, North and South Cappadocian (cf. Dawkins 1916) and an intermediate one, named Central Cappadocian (cf. Janse forthcoming)¹⁰ showing intra-dialectal divergence.

According to Janse (2001), Turkish loan verbs are completely adapted in the Cappadocian verb system. However, it is not always easy to decide how they are accommodated since, as already stated by Dawkins (1916:129), there seems to be two different forms: 1) by adding -do, -das, -da etc. or 2) -dizo which vary in the different sub-varieties of Cappadocian.

Let’s have a closer look at the data following from (5a) to (5d). We can see data from Axó in (a) and Misti in (b) -belonging to the Central Cappadocian zone- and from Aravań and Ulağaç in (c) and (d) respectively, belonging to the South Cappadocian zone¹¹.

(5)

(a) Cappadocian, Axó	<	Turkish	
/jola'dizo/ {γιολλαντίζω}	<	yollamak	
‘to fall sick (for animals)’		‘to fall sick (for animals)’	
/dayla'dizo/ {νταγλαντίζω}	<	dağlamak	
‘to cauterize’		‘to cauterize’	
/ja'fa'dizo/ {γιαῖαντίζω}	<	yaşamak	
‘to live’		‘to live’	

Data from Mavrochalyvidis (1990)

(b) Cappadocian, Misti	<	Turkish	
/γaza'dizu/ {γαζαντίζου}	<	kazanmak	
‘to earn, to profit’		‘to earn, to profit’	
/γavu'stizu/ {γαβουστιζου}	<	kavuşmak	
‘to meet’		‘to meet’	
/batur'dizo/ {μπατουρντίζω}	<	batırmak	

⁹ Cappadocian was found under Turkish influence for the first time in 11th century after the Seljuk invasion and the subsequently in the 14th century after the conquest of Asia Minor by the Ottoman Turks.

¹⁰ For a more detailed categorization of the Cappadocian varieties into zones see the Appendix.

¹¹ We should notice that in Cappadocian and Pontic the realization of /i/ in -izo, is often subject to the Turkish vowel harmony laws (-izo after a stem with /a/ or /ə/, -əzo with /o/ or /u/, -üzo with /ö/ or /ü/ and -izo after /e/ or /i/). However, its realization is often different than the Turkish vowel harmony would impose (Dawkins 1916:67) and in many cases the harmony is not observed at all and the realization of the suffix is always -izo (cf. Dawkins 1916:69, Janse forthcoming, Papadopoulou 1955).

'to waste'

'to ruin, to spoil'

Data from Kotsanidis (xx)

(c) Cappadocian, Araván< **Turkish**

/benze tizo/ {μπενζετιζω}

< bezentmek

'compare to'

'compare to'

/bile dizo/ {μπιλεντιζω}

< bilemek

'to sharpen'

'to sharpen'

/bindir dizo/ {μπιντιρντιζω}

< bindirmek

'to put sb on an animal's back'

'to embark'

Data from Fosteris & Kesisoglou (1950)

(d) Cappadocian, Ulağaç< **Turkish**

/ara du/ {αραντού}

< aramak

'to seek, to look for'

'to seek, to look for'

/biri ktu/ {μπιρικτού}

< birikmek

'to get together'

'to get together'

/ja ja du/ {γιαζαντού}

< yaşamak

'to live'

'to live'

Data from Kesisoglou (1951)

Looking at the examples above and reinterpreting these observations in terms of morphological constituents and accommodation strategies, as shown in the previous section, we could say that there is a variation in terms of accommodation mechanisms across the different sub-varieties of Cappadocian.

Axó, Araván and Misti seem to accommodate Turkish loans through the attachment of *-izo* suffix to an inflected Turkish verbal form (the third singular of the Past), i.e. through the indirect insertion strategy as shown in (6a):

(6) a.	dajad~	+ -izo	→	daja dizo (Class I verbs)
b.	dajad~	+ -o	→	daja do (Class II verbs)
	< dayan-dı-∅.Past.3.S.			'to stand'
	'to stand'			(adapted from Janse 2001:477)

Turkish verb loans in this case become part of the first verb inflectional class, which contains stems bearing stress and non systematic allomorphy for the perfective aspect forms (cf. Ralli 1988, 2005). We should notice, that *-izo* is a very productive suffix in different varieties of Greek, and it is systematically used for the accommodation of Turkish loans.

On the other hand, in Ulağaç, a different accommodation strategy seems to be active, since in that case, no affix comes into use¹². In this case, there is a clear preference for the direct insertion strategy, since Turkish verb loans accommodate directly -with no overt marker- to the second inflectional class containing stems

¹² The addition of *-o*, is categorized as a direct insertion mechanism, since inflectional suffixes are category neutral and no element marks the category of verb.

which bear no stress and show a systematic allomorphic pattern, described by (Ralli 1988, 2005) as X(a) ~ Xi¹³. Accommodation can be seen in (6b):

Even though the mechanism is different, in the former case indirect insertion, while in the latter direct insertion, as already noticed by Janse (2001) in a different vein, structural factors in both cases play crucial role to the selected strategy. More specifically, adaptation can be accounted for in terms of equivalence between the Turkish and the Greek definite past and due to the fact that the Turkish past, *dayadi* for example, coincides with the Greek perfective stem. Additionally, variation between the two schemata is triggered by the fact that this perfective allomorphic stem in *-i* can be part both of *daya'do* and *daya'dizo* as shown in (7):

(7)	dayan-di-∅	→	da'jadi-s-a
	'to stand'. Past.3S.		'to stand'.Past.3S.
	da'jadi-s-a	⇒	da'ja'd-izo
	'to stand'.Past.1S.	=>	da'ja'd-o

(Adapted from Janse 2001: 477)

It is indeed the case that, in general, Greek verbs both in *-o* and *-iz(o)* have the same allomorph, i.e. ~Xi, for the perfective aspectual value. You see in the examples that follow under (8) the underlined allomorphs of *zoγra fizo* and *aya'po*, sharing the same ~Xi pattern, i.e. *zoγrafi* and *ayapi*:

(8)	zoγra fizo	zo'γrafi-s-a	(Class I verb)
	'to paint'.1SG.PRES.	<u>zoγrafi</u> ~PERF.ASPECT.1SG.PAST.	
	aya'po	a'γapi-s-a	(Class II verb)
	'to love'.1SG.PRES	<u>ayapi</u> ~ PERF.ASPECT.1SG.PAST.	

In other words, we could say that the phonological and structural equivalence of the loan verb form with the allomorphic stem for the perfective aspectual value can account for the different patterns. However, the question arising is whether the showing preference towards a different accommodation schema in the different sub-varieties could be interpreted in terms of contact, supposing in other words that Ulağaç variety is more heavily influenced by Turkish than the other varieties just mentioned or if the selected pattern is the one favoured in structural terms.

It is true that Ulağaç and the other South Cappadocian varieties -especially the Southeast (Ulağaç and Semendere) are more heavily influenced from Turkish. According to Dawkins 1916: 209) in this zone "*the turkish element is at its strongest*". This observation could serve as a strong argument in favour of the typological hierarchy proposed by the Wichmann & Wohlgemuth since, at least so far, where the Turkish influence is said to be more intense, a different accommodation mechanism (i.e. direct and not indirect insertion) is active.

However, the situation is not really so uniform. Examples of Turkish verbal loans in other Cappadocian sub-varieties, do not seem to verify this thesis. The available data from the other Cappadocian varieties are not uniform either. Direct

¹³ Following Ralli (1988, 2005), X(a) form characterizes paradigms of an imperfective aspect (present, imperfect and future continuous paradigms), while the Xi form those of a perfective aspect (aorist and simple future paradigms).

insertion strategy is in use in some varieties belonging to the North Cappadocian subgroup which is not considered to be so heavily influenced by Turkish-, as can be seen in the examples from Anaku (9) and Floyitá (10), as opposed to Malakopi (11) and Delmesos (12) where indirect insertion, through *-izo* is active.

(9) **Anakú**

/alu'do/ {αλουντώ}	<	almak
'to take, to get'		'to take, to get'
/kapa'do/ {καπαντώ}	<	kapatmak
'to close, to shut'		'to shut'
/jarul'do/ {γιαρουλντώ}	<	yarılamak
'to split, to tear'		'to divide, to split'

(Data from Costakis 1964)

(10) **Floyitá**

/baʒər'do/ {μπαγκ[ə]ρδώ}	<	bağirmek
'to cry out'		'to cry out'
/t/əʒər'do/ {τσ[ə]γ[ə]ρντώ}	<	çağirmek
'to call'		'to call'
/kapa'to/ {καπατώ}	<	kapatmak
'to close, to shut'		'to shut'

(Data from Dawkins 1916)

(11) **Malakopi**

/baʎa'dizu/ {μπασ'λαντίζου}	<	başlamak
'to begin'		'to begin'
/jurul'dizu/ {γιουρουλντίζου}	<	yurulmak
'to be tired'		'to be tired'
/düşün'düzu/ {ντ[ü]σ[ü]ν.ντ[ü]ζου}	<	düşünmek
'to think, to consider'		'to think, to consider'

(Data from Dawkins 1916)

(12) **Delmesos**

/anla'dəzo/ {ανλαντ[ə]ζω}	<	anlamak
'to understand'		'to understand'
/ara'dəzo/ {αραντ[ə]ζω}	<	aramak
'to seek'		'to seek'
/batər'dəzo/ {μπατ[ə]ρντ[ə]ζω}	<	batırmak
'to dip'		'to dip'

(Data from Dawkins 1916)

The examined data show that a typological hierarchy of mechanisms in terms of intensity of contact and bilingualism cannot account adequately for the observed divergence and further investigation is necessary in the systems of the different sub-varieties of Cappadocian in order to account for it, something that was not possible till now since the available data are not equally ample for all the different varieties. However, it seems that further investigation of the available – productive structural schemata in each sub-variety could shed some light to the observed divergence.

3.2 Aivaliot

To broaden the picture, let us now turn to the Aivaliot dialect. Asia Minor populations speaking the Aivaliot dialect, although living in a purely Turkish environment, (in the gulf of Ayvalik and today's Cunda in West Turkey), from the end of the 16th till the beginning of the 20th century, never undergone a heavy structural pressure. On the contrary, due a Sultan's decree (at the 17th century), they enjoyed administrative and religious autonomy which provided them homogeneity and constrained the Turkish-Greek contact only on the level of commercial and administrative contact and not on the every-day one. It's worth mentioning that very few women spoke Turkish, while men used it only in trade and administration, when necessary, as opposing to Cappadocian, where bilingualism was extremely spread.

Given these, adaptation of loan verbs from Turkish in Aivaliot can be seen in the examples below:

(13)	/kaza 'dizu/ {καζαντίζου} & /kaza 'do/ {καζαντώ}	< kazanmak
'to earn, to become rich'		'to earn'
	/daja 'dizu/ {νταγιαντίζου} & /daja 'do/ {νταγιαντώ}	< dayanmak
'to stand, to sustain'		'to stand, to sustain'
	/sasir 'dizu/ {σασιρντίζου} & /sasi 'rdo/ {σασιρντώ}	< saşırmak
'to lose one's head'		'to lose one's head'
	/axtar 'dizu/ {αχταρντίζου} & /axtar 'do/ {αχταρντώ}	< aktarmak
'to turn sth upside down'		'to carry, to transfer'
	/katsir 'dizu/ {κατσιρντίζου} & /katsir 'do/ {κατσιρντώ}	< kaçırmak
'to escape, to get away'		'to escape, to get away'

(Data from the Laboratory of MGD, University of Patras)

What can be seen from the examples above is that in Aivaliot, there is systematic alternation between the two different accommodation schemata in use, e.g. *kaza'dizu* & *kaza'do* 'to earn, to become rich'. In the case of *-iz(o)* the indirect insertion mechanism is active, since there is an overt affix accommodating the loan verbal form, while in the case of *-o*, the direct insertion mechanism, where the loan verb form is adapted with no overt morphological marker to the 2nd inflectional class. In structural terms, alternation between the two schemata can be accounted for on the basis of the phonological and structural equivalence of the perfective allomorphs of *-iz(o)* and *-o* verbs, shown in (13) above (see also Ralli 2009b for a similar claim).

In terms of the Wichmann & Wohlgemuth (2008) hypothesis, alternation among different schemata should be accounted for in terms of increasing degree of bilingualism or more intense contact. But could this be the case, especially if one takes into account that Aivaliot cannot be placed so high in the hierarchy of contact as let's say Cappadocian? In our view, we cannot talk about 'strong intense contact' in terms of Thomason & Kaufman (1988), not at least as strong as in the case of Cappadocian, where agglutinative structures can be found.

On the contrary, our claim is that, alternation between the two different strategies cannot be interpreted in terms of increase in the intensity of contact. The system of the dialect offers support favoring a structural interpretation, more

specifically competition between the two processes, (affixation in the case of *-izo* and zero derivation in the case of *-o*) resulting in their alternation. It is crucial to notice that this alternation between the two processes is observed not only on the level of loan verb adaptation but on native stems as well (cf. Melissaropoulou & Ralli 2007, Melissaropoulou 2007, Ralli 2009a). In (14) below we can see instances of alternation between *-izo* and *-o* forms in native stems:

- (14) *δραστσε* 1-*izu* {δραστσελ-ίζου} & *δραστσε* 1-*o* {δραστσελ-ώ}
 ‘to lope’
ζιματ 1-*izu* {ζιματ-ίζου} & *ζιματ* 1-*o* {ζιματ-ώ}
 ‘to scald’
χλίμ 1-*dr-izu* {χλιουμντρ-ίζου} & *χλίμ* 1-*dr-o* {χλιουμντρ-ώ}
 ‘to neigh’
φρουκα 1-*izu* {φρουκαλ-ίζου} & *φρουκα* 1-*o* {φρουκαλ-ώ}
 ‘to sweep’
μυρμιδ 1-*izu* {μυρμηδ-ίζου} & *μυρμιδ* 1-*o* {μυρμηδ-ώ}
 ‘to shudder’
γκουτσί 1-*izu* {γκουτσίζου} & *γκουτσί* 1-*o* {γκουγκ-ώ}
 ‘to groan’

In several cases, as can be seen in the examples following in (15) the forms in *-o* have prevailed in Aivaliot, for example *zvo* instead of *ʒvino*, providing further evidence in favor of the increasing productivity of *-o* verbs, since according to Ralli (2009a) the leveling of verbs with various irregular allomorphic stems according to the X(a) ~ Xi pattern establishes a uniform stem-allomorphy pattern, and optimizes lexical representations by increasing their conformity to the system.

- (15) *zvo* (zv(a) ~ zvi) instead of *ʒvino* (zvin ~ zvi)
 ‘to put out, to blow out, to turn off’
fto (ft(a) ~ fti) instead of *ftino* (ftin ~ fti)
 ‘to spit’
arpo (arp(a) ~ arpi) instead of *arʒpazo* (arpaz ~ arpay)
 ‘to catch’

3.3 Pontic

Lastly, Pontic was also in a long term contact with Turkish, since it was as well spoken, already well established according to Browning 1991¹⁴, from the 12th century till the exchange of populations in 1923, in a Turkish environment, in Northwest Turkey. However, we cannot talk about heavy ‘overwhelming long-term cultural pressure’, as in the case of Cappadocian, since Pontic people were functioning for many centuries, as a closed community living on the borders, fighting with Turkish nomads and maintaining to a greater extent their homogeneity.

The vast majority of loans in all different sub-varieties of Pontic¹⁵ are accommodated via the indirect insertion strategy as well, but with preference for

¹⁴ For further details cf. Browning (1991:170-171).

¹⁵ Manolis Triandafyllides ([1938] 1981:288) proposed a tripartite categorisation of Pontic in zones: a) the zone of Oinoi, b) the zone of Trebizond and c) the zone of Chaldia, acknowledging that

another suffix, i.e. *-evo*^{16/17} (cf. Papadopoulos 1955:144). See the examples under (16) below:

- (16) /jara'ɛvo/ {γιαραεύω} < yaramak
 'to be of use, to serve' 'to be of use, to serve'
 /jola'ɛvo/ {γιολαεύω} < yollamak
 'to send, to see sb off' 'to send, to see sb off'
 /zaiɸla'ɛvo/ {ζαϊφλαεύω} < zayɸlamak
 'to slim' 'to slim'
 imzala'ɛvo {ιμζαλαεύω} < imzalamak
 'to sign' 'to sign'
- (Data from Tsopouridis 2002)

What is really interesting is that in several cases both the nominal and the verbal stem are borrowed from Turkish. See the examples under (17) below:

- (17)
 /jarala'ɛ(e)fkome/{γιαραλάυ(ε)κομαι} < yaranlamak
 'to get injured / hurt' 'to get injured / hurt'
 /jara/ {γιαρά} < yara
 'wound' 'wound'
 /imzala'ɛvo/ {ιμζαλαεύω} < imzalamak
 'to sign' 'to sign'
 /imza/ {ιμζά} < imza
 'signature' 'signature'
 /kampɸila'ɛvo {καμψιλαεύω} < kamɸılamak
 'to lash' 'to lash'
 /kampɸin/ {καμψίν} < kamɸı
 'lash, strap' 'lash, strap'

What is particularly interesting in Pontic is the fact that *-evo* is not used only in loan verb accommodation, but (it) is systematically combined with nominal bases of Turkish origin to form verbs with no verbal equivalents, like those seen under (18):

- (18)
 /xuzme'tevo/ {χουζμετεύω} < /xu'zmet/ {χουζμέτ} < hizmet
 'to serve' 'service'
 /ɸai'ɸevo/ {ɸαϊπέύω} < /ɸa'ɸp/ {ɸαϊπ} < kayɸp

this zone was more influenced from Turkish and shared common characteristics with Farasiot. Cf. Triantafyllides (1938) and Kontosopoulos (2001) for a more detailed classification of Pontic in dialectal zones and sub-varieties.

¹⁶ According to Papadopoulos (1955:144) and Tsopouridis (2002), in the areas of Kotiora and Nikopoli *-evo* is realized as *-avo*, due to coarticulation of /a/ and /e/ vowels, e.g. *zaiɸlavo* instead of *zaiɸla'ɛvo*.

¹⁷ There is a really restricted number of verbal forms suffixed with *-izo*, which have a free variant in *-evo* as well. E.g. /tae'n-izo/ {ταενίζω} & /tae'n-evo/ < dayanmak
 'to stand' 'to stand'

'to disappear'		'out of sight, away, missing'
/zabu'nevo/ {ζαμπουνεύω}	<	/za'bu'ns/ {ζαμπούν'ς} < zebun
'to be slim, stentless'		'slim, strengthless'

The most important pattern as far as *-evo*'s behavior with respect to loans is concerned, is the fact that both the nominal and the verbal loan forms are accommodated in Pontic through the use of *-evo* to form verbs. Crucially, the same base -both nominal and verbal- is found in Pontic to be attached to *-evo* suffix in order to form a verb. Taking into account the fact that loan nouns are more easily borrowed and from that viewpoint they precede verbal loans, it's not so untenable to presume that in the case of Pontic generalization of an existing schema, i.e. affixation with *-evo*, took place to accommodate verbs through the addition of *-la-*, an element which was frequent in verbal loan bases but not to nominal ones. However, there are no historical sources available one could use to test the hypothesis.

In the examples that follow in (19) we can see instances of the same stem, both nominal and verbal, attaching to *-evo*. The difference, in the two different kinds of bases is marked in Pontic, through the use of *-la-*¹⁸, a Turkish suffix which is used productively to form verbal bases from nominal ones (c.f. Kornfilt 1997:453, 455).

(19)

/suva'evo/ {σουβαεύω} & /suvaxla'evo/ {σουβαχλαεύω} <	sıva
'to plaster'	'plaster'
Note: the corresponding Turkish verb is sıvamak	
/cevezevo/ {κεβεζεύω} & /ceveze levo/ {κεβεζελεύω} <	geveze
'to chatter'	'chatty'

Note: the corresponding Turkish verb is gevezelik etmek

In some cases this *-la-* suffix, which is recognized as a marker accommodating verbal stems, expands a) to other Turkish loans which are not verbal, which do not have verbal equivalents with *-la-*, (20a) and b) crucially to Greek bases as well (20b), as can be seen in the examples below:

(20)

a) /pæli la'evo/ {πάλληλαεύω} <	/pæ'lis/ {πάλλης} < beli
'to imprint, to stamp'	'clean, evident' 'clean, evident'
/rezilæ'efkume/ {ρεζιλᾶεύκουμαι} <	/rezil/ {ρεζίλ} < rezil
'be held up to ridicule'	'ridicule' 'indecent, shameless'
b) /cenurla'evo/ {καινουρ-λα-εύω} <	/cenuræ/ {καινούρα}
'to renew'	'new'
/sirala'evo/ {σειραλαεύω} <	/si'ra/ {σειρά}
'line up, arrange in rows'	'row'
/nani la'evo/ {νανιλαεύω} <	/nani/ {νάνι}
'to lulle'	'sleep'

¹⁸ Cf. Dawkins (1916:130) and Janse (forthcoming) for some marginal examples with *-lan-* Turkish suffix marking reflexiveness in Cappadocian.

These data offer a clear counterexample to the general claim made by Moravcsik (1975, 1978: 111-112) that, if verbs are borrowed, they seem to be borrowed as if they were nouns, or to its moderate version of underspecified insertion by Wichmann & Wohlgemuth (2008). On the contrary, Pontic seems to expand the same mechanism used for the verbalization of Turkish nominal bases to the accommodation of verbs, i.e. the Turkish verbal marker *-la-*.

Lastly, we should notice that the use of *-evo* suffix is found in the other varieties in study as well. Both in Cappadocian and Aivaliot, nominal bases of Turkish origin -fully adapted in the dialectal system- form verbs when attached to *-evu /evyu (<-evo)* suffix. Examples can be seen under (21) below:

(21) Aivaliot

/zabu ʔevyu/ {ζαμπουν-εύγου}	<	/za buʔs/ {ζαμπούν'ς}	<	zebun
'to fall sick'		'sick'		
/bata ʔevyu/ {μπαταλεύγου}	<	/ba ta ʔs/ {μπατάλ'ς}	<	battal
'to disuse, to destroy'		'of no use'		
/xadzi ʔevyu/ {χατζιρεύγου}	<	/xa ʔir/ {χαζίρ}	<	hazır
'to prepare, to get ready'		'ready'		
/murda ʔevyu/ {μουρνταρ-εύγου}	<	/mur ʔdaris/ {μουρντάρης}	<	murdar
'to be/to get dirty'		'dirty'		

(22) Cappadocian (Axó, Araván, Misti, Ulağaç)

/xuzu ʔevu/ {χουζουρ-εύου}	<	huzur {χουζούρ}	<	huzur
'to have a long lie-in'		'lie-in'		
/saka ʔevu/ {σακατ-εύου}	<	sakatis	<	sakat
'to cripple'		'cripple'		
/xardzi ʔevu/ {χαρτζι-εύου}	<	xardzi	<	harsi
'to spend'		'expense'		

In our view, this choice is not accidental but is due to the fact that *-evo* in Greek is in general very productively combined with nominal bases to form verbs.

4. Conclusions

To conclude, we would say that we have examined loan verb accommodation mechanisms from Turkish in three different varieties of Greek. Dialectal data show variation between the mechanisms of direct and indirect insertion. This variation, which according to Wichmann & Wohlgemuth (2008) typological classification of loan verb accommodation strategies could be associated with sociolinguistic factors (i.e. intensity of contact and a higher degree of bilingualism) cannot, solely, account adequately neither for the cross-dialectal nor for the intra-dialectal divergence.

The sub-varieties of Cappadocian, which are considered as an instance of 'overwhelming long-term cultural pressure' (Thomason & Kaufman 1988:50) show a diverging behaviour. The Aivaliot dialect, which cannot be placed as high as Cappadocian in terms of Thomason and Kaufman scaling of intensity of contact, exhibits variation between the two strategies. In Pontic, which in terms of contact

could be placed in an intermediate position, not so high as Cappadocian, but definitely higher than Aivaliot, indirect insertion strategy is in use, with a different suffix, i.e. *-evo* which is used productively not only for the accommodation of verbs but for the verbalization of nominal bases of Turkish origin as well, often with the addition of *-la-* affix.

Without ignoring the crucial role of the intensity of contact and the degree of bilingualism for the adoption of loan verbs, we claimed that structural factors, i.e., phonological and structural equivalences, the notion of productivity and the competition among the existing processes could account both for the preferred mechanism and the alternation between different strategies – schemata.

More particularly, we have seen that alternation between the two different mechanisms cannot be interpreted as a case of heavier pressure in Aivaliot from Turkish, but proves to be a generalized schema involving competition between zero derivation and affixation with *-izo*, triggered by a strong tendency towards the establishment of a uniform stem-allomorphy pattern aiming at the optimization of lexical representations Ralli (1988, 2009a). Surprisingly, this alternation is not found -at least systematically- in Cappadocian, a phenomenon which merits further investigation, with enrichment of data from all different sub-varieties. Lastly, in Pontic no alternation of schemata is observed. In this case, the dialectal data offer an extra counterexample to the thesis that loan verbs are entering the system of the target language as nouns or underspecified (Moravcsik 1975, 1978, 2003) since, in Pontic the same suffix, i.e. *-evo* with the addition of *-la-* is used to mark verbal loans, while in the other varieties, a different suffix is used to verbalize nominal bases of Turkish origin.

However, we should say that this study is only in the beginning. More systematic research and enrichment of data, both from Turkish and from other source languages is needed in order to test where there is variation when a) the typological features of donor and / or recipient language change and b) when the language contact situation changes. In our long-term goals are a cross-dialectal typology of verbal borrowing patterns as well as a typological hierarchy of social, grammatical and lexical factors affecting the borrowability of verbs. What is proven though, at least so far, is within the spirit of Singh thesis that, although history decides the change, it is the grammar that will decide which road it will take (Singh, 1996 2008).

5. References

- Anastasiadi- Symeonidi, A. (1994). *Neological Borrowing in Modern Greek* [Νεολογικός Δανεισμός της Νεοελληνικής]. Thessaloniki.
- Andriotis, N.P. (1948). *The linguistic variety of Farasa* [Το γλωσσικό ιδίωμα των Φαράσων]. Athens: Ikaros.
- Bakker, Peter. (2005). Intertwining and Michif. Paper presented at the Romanisation Worldwide Conference. Bremen, 5-9 May 2005.
- Browning, R. (1991). *Medieval and Modern Greek* [Η ελληνική γλώσσα, μεσαιωνική και νέα]. Greek translation M. N. Konomi. Athens: Papadimas.
- Costakis, A. (1964). *Le Parler Grec d'Anakou*. Athènes: Centre d'Études d'Asie Mineure.
- Dawkins, R. (1916). *Modern Greek in Asia Minor*. Cambridge: Cambridge University Press.
- Drettas, G. (1999). The Greek –Pontic dialectal group [Το ελληνο-ποντιακό διαλεκτικό σύνολο]. In A.-F. Christides (in collaboration with M. Arapopoulou and G. Giannoulou) (ed.), *Dialectal enclaves of the Greek Language* [Διαλεκτικοί θύλακοι της ελληνικής γλώσσας]. Athens: YPEPTH & Center for the Greek Language
- Drettas, G. (2000). Marques de *focus* en grec commun et en pontique. *Studies in Greek Linguistics* 1999, 122-131.

- Field, F. (2002). *Linguistic Borrowing in Bilingual Contexts*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Fosteris, D. & Kesisoglou, I.I. (1950). *Vocabulary of Aravani* [Λεξιλόγιο του Αραβανί]. Athens: French Institute of Athens.
- Haspelmath, M. (2008). Loanword typology: Steps toward a systematic cross-linguistic study of lexical borrowability. In T. Stolz, D. Bakker & R. Salas Palomo (eds.) *Aspects of language contact: New theoretical, methodological and empirical findings with special focus on Romancisation processes*. Berlin/New York: Mouton de Gruyter, 43-62.
- Haugen, E. (1950). The analysis of linguistic borrowing. *Language* 26: 210 – 231.
- Hock, H. & B. Joseph. (1996). *Language History, language Change and Language Relationship*. Berlin: Mouton de Gruyter.
- Jäger, A. (2004). The cross-linguistic function of obligatory ‘do’-periphrasis. Paper presented at the 2004 Conference of the Australian Linguistic Society. Available at <http://66.249.93.104/search?q=cache:xcVAbVnhYrEJ:dspace.library.usyd.edu.au:8080/bitstream/123456789/111/1/ALS-20050630-AJ.pdf+loan+verbs&hl=en>. Accessed at 2010-04-13, 5:59 pm.
- Janse, M. (2001). Morphological Borrowing in Asia Minor. In G. Aggouraki, A. Arvaniti, J.I.M. Davy, D. Goutsos, M. Karyolaimou, A. Panagiotou, A. Papapavlou, P. Pavlou, A. Roussou (eds.), *Proceedings of the 4th International Conference on Greek Linguistics*. (Nicosia, 17-19 September 1999). Thessaloniki: University Studio Press, 473-479.
- Janse, M. (forthcoming). “Cappadocian”. In Tzitzilis (ed.) *The Greek language and its dialects* [Η ελληνική γλώσσα και οι διάλεκτοί της]. Thessaloniki.
- Kaltsa M. & Sitaridou I. (2010). Topicalization in Pontic Greek. *Paper Presented at the 4th International Conference of Modern Greek Dialects and Linguistic Theory* (Chios, June 11-14 2009).
- Kesisoglou, I.I. (1951). *The dialectal variety of Ulağaz* [Το γλωσσικό ιδίωμα του Ουλαγάτζ]. Athens: French Institute of Athens.
- Kontosopoulos, N. (2006). *Modern Greek Dialects and Idioms* [Διάλεκτοι και ιδιώματα της νέας ελληνικής]. 4th Edition. Athens: Grigori Editions.
- Kornfilt, J. (1997). *Turkish. Descriptive Grammars*. Routledge: London and New York.
- Kotsanidis, L. (xx). *The dialect of Misti Cappadocia* [Το γλωσσικό ιδίωμα του Μιστί Καππαδοκίας]. Kilkis: Gnomi Kilkis-Paionias.
- Mackridge, P. (1999). Grekophony in the area of Ofi (Pontus) [Η ελληνοφωνία στην περιοχή του Όφι (Πόντος)]. In A.-F. Christides (in collaboration with M. Araproulou and G. Giannoulorou) (ed.), *Dialectal islands of the Greek Language* [Διαλεκτικοί θύλακοι της ελληνικής γλώσσας]. Athens: YPEPTH & Center for the Greek Language, 25-30.
- MacMachon, A. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Mavrochalyvidis, G. (1990). *Axo Cappadocia* [Η Αξό Καππαδοκίας]. Monography 2nd volume. Athens: O., P. & A. Mavrochalyvidi.
- Melissaropoulou, D. (2007). *Word Formation in the Asia Minor Dialect of Kydonies and Moschonisia (Aivaliot)* [Μορφολογική περιγραφή και ανάλυση του Μικρασιατικού ιδιώματος Κυδωνιών και Μοσχονησίων: η παραγωγή λέξεων] Unpublished PhD Dissertation. University of Patras.
- Melissaropoulou, D., Ralli, A. (2007). The verbal system of the Asia Minor Dialect of Kydonies and Mosxonisia [Το ρηματικό σύστημα της διαλέκτου των Κυδωνιών και Μοσχονησίων.]. In P. Kitromilides and P. Michailaris (eds.) *Proceedings of the fifth International History Conference “Mitilene and Aibali: A mutual dealing in North-Eastern Aegean Sea”* (6-9 October 2003, Mytilene). Athens: National Research Institute, 203-211.
- Michelioudakis, D. & Sitaridou, I. (2010). Syntactic micro-variation in Pontic: Dative constructions. *Paper Presented at the 4th International Conference of Modern Greek Dialects and Linguistic Theory* (Chios, June 11-14 2009).
- Muysken, P. (2000). *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, C. (2002). *Language Contact: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Moravcsik, E. (1975). Borrowed verbs. *Wiener Linguistische Gazette* 8, 3-30.
- Moravcsik, E. (1978). Universals of Language Contact. In J. H. Greenberg et al. (eds.), *Universals of human language, vol. 1 Method and Theory*. Stanford, CA: Stanford University Press, 93-122.
- Moravcsik, E. (2003). “Borrowed Verbs”. Ms.

- Mougeon, R. and E. Beniak. (1989). Language contraction and linguistic change: The case of Welland French. In N. C. Dorian (ed.) *Investigating obsolescence: Studies in language contraction and death*. Cambridge: Cambridge University Press, 287-312.
- Myers-Scott, C. (2002). *Language Contact: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Papadopoulos, A. (1955). *Historical Grammar of the Pontic Dialect [Ιστορική γραμματική της ποντικής διαλέκτου]*. Pontic Archive, Appendix 1. Athens: Committee for Pontian Pontic Studies.
- Poplack, S., Sankoff, D. and C. Miller (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47-104.
- Ralli, A. (1988). *Eléments de la morphologie du grec moderne: La Structure du Verbe*. Unpublished Ph.D. thesis, University of Montreal.
- Ralli, A. (1995). Structure et forme des emprunts français dans la langue grecque. *Europe Plurilingue* 7, 77-85.
- Ralli, A. (2005). *Morphology [Μορφολογία]*. Athens: Patakis.
- Ralli, A. (2009a). "Morphology meets Dialectology: insights from Modern Greek dialects". *Morphology* 19, 87-105.
- Ralli, A. (2009b). "Morphology in Language Contact: Verbal Loanblend formation in Aivaliot". Paper presented at the Conference *Morphologies in Contact*, Bremen, October, 1-3 2009.
- Singh R. (1996). *Linguistics Theory, Language Contact and Modern Hindustani*. New York: Peter Lang.
- Singh, R. (2008). 'Language Contact'. Lecture given at the Master Program "Comparative Linguistics and Language Variation" Department of Philology, University of Patras, May, the 15th 2008.
- Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.
- Triantafyllides, M. [1938] (1981). *Completed Work*. 3rd vol., *Modern Greek Grammar. Historical Introduction*. Thessaloniki: Institute of Modern Greek Studies [Manolis Triantafyllides Foundation].
- Tsopouridis, T. (2002²). *Lexicon of the Pontic Dialect [Λεξικό Ποντιακής Διαλέκτου]*. 2nd volume, from Pontic to Modern Greek (1st edition 1998). Θεσσαλονίκη: Tsopouridis editions.
- Weinreich, U. (1968²). *Languages in Contact* (1st edition 1953). The Hague: Mouton.
- Whitney, W. (1881). On mixture in language. *Transactions of the American Philosophical Association*, 12:1-26.
- Wichmann, S., Wohlgemuth, J. (2008). Loan verbs in a typological perspective. In T. Stolz, D. Bakker & R. Salas Palomo (eds.), *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Berlin/New York: Mouton de Gruyter, 89-121.
- Van Hout, R., P. Muysken. (1994). Modeling lexical borrowability. *Language Variation and Change*, 6 (1): 39-62.

Appendix

Subgroupings of Cappadocian

North Cappadocian

- Northwest Cappadocian
 - ✓ Sılata
 - ✓ Anakú
 - ✓ Floyitá
 - ✓ Malakopı
- Northeast Cappadocian
 - ✓ Sinasós
 - ✓ Potámya
 - ✓ Delmesó

Central Cappadocian

*Loan verb adaptation in Greek dialectal variation:
A first approach*

✓ Axó

✓ Mistĩ

South Cappadocian

➤ Southwest Cappadocian

✓ Araván

✓ Ferték

➤ Southeast Cappadocian

✓ Ulağáç

✓ Semenderé

(From Janse forthcoming)

Topicalisation in Pontic Greek

Maria Kaltsa
Aristotle University of Thessaloniki
mkaltsa@enl.auth.gr

Ioanna Sitaridou
University of Cambridge, Queens' College
is269@cam.ac.uk

1. Introduction

Pontic Greek is a variety of Greek which was historically spoken outside the area which now constitutes the Greek state. Today, as a consequence of the Treaty of Lausanne (1923), it is spoken both within and outside Greece. Within Greece it is mainly spoken in Macedonia, Thrace, and to a lesser extent in Attica. Outside Greece it is spoken in the Pontus region –the historical berceau of all Pontic varieties– but also in Istanbul, in Caucasus and by diaspora communities across the world. Although Pontic in Greece seems to be robust in terms of number of speakers, in real terms the majority of speakers is severely attrited. Indicative of the attrition situation is that although the number of Pontians is quite significant (above 2 million in Greece alone) only a fraction of the population (200,000 or 300,000 depending on the estimates) is reported to be active speakers of the dialect. Due to the geographical dispersion of Pontic, it is important to note that the term Pontic, synchronically, can only be used as an “umbrella” term for the various subdialects, which, crucially, can diverge significantly from each other (e.g., “Christian” vs. “Muslim” Pontic, cf. Mackridge 1987). For the purpose of this paper we use the term “Pontic” to refer to the Pontic varieties of Northern Greece.

Greek dialectal syntax is notoriously understudied primarily because of all the efforts –perpetuating at both social and institutional level– to erase dialectal variation and instead, impose linguistic uniformity in the name of Standard Modern Greek (henceforth SMG) (for the same view see also Ralli 2007). Within this context, work on dialectal syntax is urgently needed and our present article aims at contributing towards this direction. The goal of the article is twofold: first, to describe the discourse phenomenon of topicalisation in Pontic syntactically; and, second, to suggest a (cartographic) analysis casted within the generative framework thus making the present work the first attempt of this kind.

The paper is organised as follows: Section 2 presents the main syntactic features of Pontic that differentiate it from the standard variety. Section 3 describes our methodology. Section 4 discusses topicalisation strategies in Pontic. Section 5 proposes a syntactic analysis of topicalisation in Pontic. Finally, section 6 concludes the discussion.

2. Syntactic variation in Pontic Greek and SMG

Although the Pontic variety spoken in Greece is by far the best described Greek dialect (cf. Oikonomidis 1958, Papadopoulou 1955, Tombaidis 1988, 1996, Drettas 1997, *inter alios*), still little is known about its syntax. For this reason, in the current section we identify the main syntactic features of Pontic (pertaining to the structure of the DP, vP, CP) which could be used as diagnostics for determining the syntactic isoglosses between Pontic and SMG as well as among Greek dialects, in general.

First, let us consider the most well-studied syntactic phenomenon of Pontic namely, the distribution of clitics (cf. Pappas 2006, Revithiadou 2008) which alone, according to Condoravdi & Kiparsky (2001:1-3), is a sufficient criterion for a taxonomy of the Greek varieties. Pontic clitics are strictly enclitics, as shown in (1):

- (1) Edoken to jon ats ton Lazaron ... ke ipen aton (Pontic)
give-3SG-PAST the son her the Lazaros-ACC ...
and say-3SG-PAST he-acc

‘She gave her son, Lazarus ... and told him’
(Kyriakidis 1998: 30)

Second, although the main distributional rule in DP constructions is that the qualifying element always precedes the qualified (cf. Drettas 1997:183, Janse 2002:221) as in SMG, Pontic differs in having obligatory determiner spreading whereas in SMG determiner spreading is optional (2b) and obligatory (2c') only when the adjective is postposed:

- | | |
|--|--------------|
| (2) a. o tranon o ðeskalon
the big the teacher
(Tombaidis 1988:61) | (Pontic) |
| b. o meyalos (o) ðaskalos
the big (the) teacher | (SMG) |
| c. *o ðaskalos meyalos
the teacher big | (SMG/Pontic) |
| c'. o ðaskalos o meyalos
the big (the) teacher
'the big teacher' | |

Additionally, although adjectival possessives are postposed in both SMG and Pontic, it is only in the latter that we find possessive spreading by means of a purely affixal possessive (3a):

- | | |
|--|----------|
| (3) a. to kalom to peðim
the good-POSS the child-POSS
(Janse 2002: 222 & Drettas 1997:166) | (Pontic) |
| b. *to kalo mu to pedi mu
the good POSS the child POSS
'My good child' | (SMG) |

Third, in dative constructions in Pontic (see also Drettas 1997, Tombaidis 1996, Michelioudakis and Sitaridou, this volume) the 'inherent' vs. 'structural' distinction is possibly collapsed into the latter. Objects, regardless of whether they are direct or indirect, are always in the accusative (4):

- | | |
|---|----------|
| (4) Ipa ton peðan tin aliθian
said-1SG the boy-acc the truth-acc
'I said the truth to the boy.' | (Pontic) |
|---|----------|

Fourth, Pontic exhibits significant variation from SMG in terms of hypotaxis since it uses more paratactic constructions. Consider, for instance, verb serialisation (cf. Drettas 1997) in (5a), which, crucially, is not found in SMG (5b) (but see Joseph (1990) on the *ela pame*, "let's go" construction). In particular, in (5a) the movement verb is paratactically connected to the second verb without any complementiser mediating.

- | | |
|---|----------|
| (5) a. as payo elepo
part-1SG-Pres see-1SG-Imperf
'I shall go and see.'
(Papadopoulos 1955: 163-164) | (Pontic) |
| b. as pao na do
part go-1SG-Pres part see-1SG-Perf
'I shall go and see.' | (SMG) |

Fifth, Pontic allows multiple *wh*-fronting, as shown in (6a), contrary to what is possible in SMG (6b):

- (6) a. *Tinan pion ospit eðiksises?* (Pontic)
 who-ACC which house show-2SG-PAST
 b. **Se pion eðikses pio spiti?* (SMG)
 to-who-ACC show-2SG-PAST which home
 ‘Which house did you show to whom?’

So far we have seen (cf. examples (1) to (6)) that significant syntactic differences are attested in all syntactic areas between Pontic and SMG. Moreover, Pontic has an additional feature namely, discourse particles which, crucially, SMG lacks. The discourse particles are identified as the clause-typing and the interrogation ones (cf. Drettas 1997). The first one to consider is the clause typing particle *kja*¹ which is etymologically related to *ke* (meaning “and”) and *ara* (meaning “consequently”). It is used primarily in dialogues and is rarely found in narratives (cf. Drettas 1997:407). The discourse function encoded by this particle is that of assertion (7):

- (7) *kja vevea ða ifes* (Pontic)
 Ass-PART certainly Fut-PART have-2SG
 ‘You will certainly have it.’
 (Drettas 1997: 408)

Another clause typing particle is *ja*.² Its discourse content is that of assertion as well. It is syntactically incompatible with interrogatives and it is positioned clause-finally immediately before an extended pause (Drettas 1997:409) (8):

- (8) *eð ke ton jeronats ja* (Pontic)
 have-3SG-Pres and the-ACC old-man-ACC+Poss-3SG-Fem Ass-PART
 ‘She also has her old man.’
 (Drettas 1997: 409)

Moving on now to the interrogation particles, we can identify at least three particles used in question clauses: *paðkim* (or *paðkimto*), *jam* and *kjam* each conveying quite distinct discourse roles. *Paðkim* (or *paðkimto*) is positioned clause-initially and functions as an intensifier which asserts certainty (Drettas 1997:411) (9):

- (9) *ta trayoðias ta kala paðkimto in ðramena?* (Pontic)
 the-ACC songs-ACC the-ACC nice-ACC Inter-PART be-3PL-Pres writ-ten-Adj-ACC
 ‘Does somebody find nice songs written?’
 (Drettas 1997: 355)

On the other hand, *jam* has the exact opposite intensive usage. It conveys either uncertainty or probability (Drettas 1997: 413-414). As (10) illustrates, both pragmatic notions are encoded via the *jam* particle and the exact discourse value can be determined

¹ Consider *ke* in SMG:

(i) *ke vevea ða erthis*
 And of course Fut-PART come-2SG
 ‘You will certainly come’

² Consider *ja* in Northern Greek:

(i). *etsi ine ja*
 that-is-how be-3SG-Pres PART
 ‘That is how it is!’

only through context:

- (10) Probability
 efoyuntane jam pernats [...](Pontic)
 be-afraid-3PL-Past Inter-PART take-3PL-Pres+Obj-3PL
 ‘They feared that they might not be taken.’
 (Drettas 1997: 413, 348)

This brief discussion on particles shows how extensively this device is used in Pontic. It is therefore, not curious that particles are also involved in marking other discourse functions, such as topicalisation and focalisation (although the latter is not discussed here, cf. Kaltsa and Sitaridou, to appear), as we shall see in section 4.

3. Methodology of data collection

The original methodological aim was to use only native data collected through elicitation and grammaticality judgement tasks. For this purpose, we selected two speakers of Pontic from Thessaloniki as our informants. We used the following criteria to select them: age (both +60); degree of exposure to Pontic (both exposed to Pontic from birth); use of Pontic in everyday life: (one on an everyday basis, the other less often); education (none with a university education albeit one with higher education); mobility (both non-mobile); language profile (no other languages apart from SMG); community status (one is considered by the community as a very able speaker); social class (both middle class). We run one-to-one pilot interviews which comprised: a) a 50-item questionnaire examining subject and object focus (Kaltsa 2007); we administered it orally so that the speakers not be confronted with the written language which may, in turn, trigger grammaticality judgements influenced by SMG given the affinity of the written medium with the standard variety; and b) free theme/narration of a story.

The pilot study showed important problems. First, the informants used excessive clitic doubling which is a very frequent in SMG, but much less so in Pontic. (11) provides further proof of a transfer from the SMG since we observe proclisis whereas we know that Pontic exhibits enclisis across the board:

- (11) to vivlio to eðavesa to olon (Pontic)
 the-ACC book-ACC the-ACC read-1SG-Past the-ACC whole-ACC
 ‘I read the whole book.’

Second, the informants used no particles in the grammaticality tasks, but only in the narration task (12):

- (12) ato emas-pa θ’etroen
 (Pontic)
 that-ACC us Top-PART Fut-PART eat-3SG-Past
 ‘It would eat us.’

Both their linguistic performance as well as their metalinguistic judgments provided evidence for the fact that the Pontic speakers today are seriously attrited and the dialect is possibly endangered despite Ethnologue’s (<http://www.ethnologue.com/>) figures asserting the opposite. However, further research is needed to consolidate this claim.

These findings led to a redesign of our main study since we could no longer rely on unattrited, robust grammaticality judgments which were not influenced by the bi-dialectalism of our informants. Therefore, we decided that our main study should mainly involve soliciting data from the written record of Pontic. For this purpose we selected texts which fulfilled the following criteria: (i) contained dialogues; (ii) publication date; (iii) availability of translation into SMG to avoid variability of interpretations. On the basis

of these criteria –to the extent that such a tall bill can be satisfied– the following texts have been employed for the identification and description of information structure in Pontic: a theatrical play dated from 1972 (Adreadis 1990); a short story dated from 1951 (Melanofrydis 2001); a selection of folktales dating from 1928 (Tombaidis 1988); and the narratives included in the grammar of Drettas (1997:515-671). After careful examination of these texts, 231 tokens encoding topicalisation have been identified (Kaltsa 2007). These data form the core of our analysis and the most representative ones in the next two sections. In analysing the data, we controlled for the following properties: sentential position; the nature of the elements that undergo topicalisation; and, the possibility of the contrastive vs. informational reading. Once the data were coded they have been further checked against the grammars of Papadopoulos (1955), Tombaidis (1988 & 1998) and Drettas (1997).

4. Topicalisation Strategies in Pontic Greek

In the literature, information structure is defined as the encoding of discourse information of an utterance through operations such as topicalisation and focalisation. Topic has been primarily identified as what the utterance is about at the level of a sentence; to put it differently, topic is the “notional subject” (at least, according to Kiss 1995:7). Meanwhile, at the level of a dialogue the topic is identified as the element that is discourse-old, and consequently known to both interlocutors. Further interpretive distinctions can be made with regards to topics: *Aboutness Topic* (ATop), which is the constituent representing the theme of the predication, namely, “what the sentence is about” (cf. Reinhart 1981, Lambrecht 1994, Frascarelli and Hinterhölzl 2007); *Contrastive Topic* (cf. Frascarelli & Hinterhölzl 2007); and *Familiar Topic* (*ibid*). In our study we utilise this tripartite distinction and focus on the former two.

There are two main strategies for topicalisation in Pontic: (i) Clitic Left Dislocation (CLLD) (with clitic doubling, henceforth CD) as in SMG, and (ii) usage of particles unlike SMG. Crucially, these two topicalisation strategies are neither interchangeable nor pragmatically identical: The former is claimed to be conveying “aboutness” whereas the latter is claimed to be conveying “contrastiveness”; (iii) there is a third strategy which is more marginal and entails a clitic-doubled *pa*-phrase –this last one is assumed to encode a discourse reading somewhere between the two aforementioned ones. In this paper we will focus on (ii) leaving aside the discussion of other topicalisation strategies (for a detailed account of the encoding of information structure, cf. Kaltsa & Sitaridou, to appear.)

It has already been noted in the literature that the use of the particle *pa* is an extremely frequent topicalisation strategy (cf. Setatos 1994, Drettas 2000, 1997). The (invariable) particle *pa* carries no stress (and consequently it is never sentence-initial), and is attached to the end of the topicalised constituent. Crucially, particle use for the encoding of discourse information is never attested in SMG. *Pa*, is etymologically related to the Ancient Greek adverb *palin*, meaning “again”, as suggested by Papadopoulos (1958-1961:3.130)³. This etymological explanation is further supported by the finding of the use of *-pal* in Cappadocian (Dawkins 1916: 631 in Janse 2002) (13):

- (13) a. ekinos-pa efxaristiθike pola (Pontic)
 he-NOM Top-PART be-pleased-3SG-Past a lot
 ‘He was deeply pleased.’
 (Tombaidis 1988: 106)
- b. k-eto-pali ... ekutθis-to piken (Cappadocian)
 and this-ACC Top-PART ox-driver-NOM+the-ACC do-3SG-Past
 ‘And this...it was the ox-driver who did it.’
 (Dawkins 1916: 424, 426)

³ The use of *pa* as a discourse marker seems to be a clear case of grammaticalisation from an adverb (lexical) to a topic marker (functional).

The data which we will discuss stem from our corpus. We noted all instances of *pa* and we controlled for the type of constituent *pa* attaches to, as well as the position of the *pa*-phrase vis-à-vis the verb. Our findings are summarised in Table 1.

Table 15: *pa*-phrases in our corpus

<i>pa</i> -attachment (231 items)	OV		VO		SV		VS	
Object	56 items	96,5%	2 items	3,5%				
Subject					120 items	97,5%	3 items	2,5%

With regards to the types of constituents *pa* attaches to let us first, let us consider instances of *pa* attaching to a subject constituent. *Pa* appears attached to pronominal subjects (14a), lexical subjects (14b), in subject DPs with DS (14c), and without being blocked by any definiteness (14d & 14e) or quantifier restrictions (14f).

(14) Subject Topicalisation (Pontic)

a. Pronoun

ego-*pa* eθaresa emen ekuikses

I Top-PART be-encouraged-1SG-Past me-ACC listen-3SG-Past

'I was encouraged that you listened to me.'

(Adreades 1990: 84)

b. Lexical DP

i popaðja-*pa* 's so mantrin ixen ðulian

the-NOM priest's-wife-NOM Top-PART to-the-ACC pen-ACC have-3SG-Past work-ACC

'The priest's wife had work at the pen.'

(Adreades 1990: 19)

c. Possessive/demonstrative (but not definite article)⁴ (with Determiner Spreading)

t'emeteron-*pa* to tixeron aikon eton

the-NOM mine-NOM Top-PART the-NOM fate-NOM of-this-kind-NOM be-3SG-PAST

'My fate was of this kind.'

(Adreadis 1990: 58)

d. Definite Subject

tin Leila-*pa* ipen na fori ta kala ta lomatats

the-ACC Leila-ACC Top-PART say-3SG-Past Mod-PART wear-3SG the-ACC nice-

ACC the-ACC clothes-ACC Poss-GEN

'S/he said to Leila to wear her nice clothes.'

(Melanofrydis 2001:33)

⁴ It may be claimed that the following example constitutes an exception to this distributional restriction since the particle attaches to the indefinite pronoun and not the entire DP:

(i) enan-*pa* litanian eftaynaton atora ta enteka ta enteka t-avyusti

a Top-PART litany-ACC do-3PL-Pres+Obj-3SG-ACC now the-ACC eleven-ACC the-ACC

eleven-ACC the-GEN August-GEN

'Recently, the 11th August, a litany took place.' (Drettas 1997: 440 ex.101)

This is however, an instance of a split-DP, see discussion on p. 12.

- e. Indefinite Subject
 o apoθamenon-pa ex topo
 the-NOM dead-NOM Top-PART have-3SG-Pres place-ACC
 ‘A dead person has its place.’
 (Adreadis 1990: 36)
- f. Quantifier subject
 ul’-pa etimanan aton
 everyone-NOM Top-PART honour-3PL-Past he-ACC
 ‘Everyone honoured him.’
 (Melanofrydis 2001: 25)

Second, consider instances of *pa*-attachment to an object (15). ‘*Pa*’ appears attached to definite objects (15a & 15b), indefinite objects (15b), *wh*-objects (15c), and objects bearing a possessive (15d).

(15) Object Topicalisation (Pontic)

- a. Definite object
 Tin aderfis’ pa m’ ayliyoris
 the-ACC sister-Poss Top-PART not forget-2SG
 ‘Do not forget your sister’
 (Melanofrydis 2001: 13)
- b. Indefinite object
 Enan-pa litanian eftaynaton atora
 one-ACC Top-PART make-3PL-Past-him now
 ‘They made a litany in his honour’
 (Drettas 1997:440)
- c. Polarity object
 Tiden pa leis
 nothing Top-PART say-2SG
 ‘Don’t say anything’
 (Andreadis 1990:45)
- d. Possessive object
 T’emon pa kap na ayrika nuniz’ aton
 the mine-PART somewhere to understand think him-ACC
 ‘Mine(mother-in-law) as soon as she felt I was this thinking of him’
 (Andreadis 1990:12)

Third, consider instances of *pa*-attachment to an adverbial (16). The *pa*-adverbials appear predominantly in the preverbal position and can be time, location or manner ones. A *pa*-adverbial can be an adverb (16c) or prepositional phrase/DP (16b).

(16) Adverbial topicalisation (Pontic)

- a. akaθarton ekino i lefkaða | atora-pa leyato ke neræskume
 unclean-NOM this-Deict-NOM the-NOM Lefkada-NOM now Top-PART say-1SG-
 Pres+Obj-3SG-ACC and make-sb-sick-1SG-Pres
 ‘Lefkada was that dirty that even now it makes me sick.’
 (Drettas 1997: 442 ex. 105)
- b. enan imeran-pa erθen enas psaras kuizmas yariðes yariðes emis ol
 exparayamen etoplaeftam ekeka
 one-ACC day-ACC Top-PART come-3SG-Past one-NOM fisherman-NOM cry-3SG-
 Pres+Obj-1PL-ACC woman-PL-ACC woman-PL-ACC we-NOM everyone-NOM get-scared-
 1PL-Past gather-1PL-Past there

‘One day a fisherman came and called the women. All of us got scared and gathered around.’

(Drettas 1997: 442 ex.107)

c. Kj- aets-pa⁵ eperenaten opis k- epantreftan

And thus-Top-PART take-3SG-PAST-her back and marry-3PL-PAST

‘And thus, he took her back and they married.’

(Drettas 1997: 448 ex.115)

Let us now discuss the position of the *pa*-phrases vis-à-vis the verb. The position of the *pa*-constituent is predominantly in the pre-verbal position and clause-initially, as shown in Table 1. The post-verbal instances are very limited and only one instance (1 occurrence out of 231 tokens) has been found in our data (17):

(17) Post-Verbal *pa*-topic

(Pontic)

lejme pios ef paraðas? Leo k eyo pa ...ekinos ekserætsen ults ekser kata onoman ke kata jenean

say-3SG-Pres+Obj-1SG-ACC who-NOM have-3SG-Pres money-ACC?

say-1SG-Pres and-EMPH I-NOM Top-PART...he-Deictic-NOM know-3SG-Pres+Obj-3PL-ACC everyone-ACC know-3SG-Pres by name-ACC and by generation-ACC

‘S/he says: Who has money?’

And I say: Everyone knows him by name and generation.’

(Drettas 1997: 551 ex. 117)

As we have seen *pa* always appears after the constituent it modifies apart from the split-DP examples where it appears interpolating between the adjective/possessive and the noun, as shown in (18):

(18) a. ðio ospita-pa ixame so xorionemun turkant

(Pontic)

two house-ACC-PL-Neut Top-PART have-1PL-Past to-the-ACC village-ACC-SG-Neut+Poss-1PL Turkish-NOM/GEN-PL-Masc

‘We had two Turkish houses at our village.’

(Drettas 1997: 438 ex. 98)

b. ekino-pa to kaimeno kaðete olen tin imeran

that-Deict-NOM Top-PART the-NOM poor-NOM sit-3SG-Pres whole-ACC the-ACC day-ACC

‘The poor thing sits all day long.’

(Adreades 1990: 27)

Examples such as (18), may *prima facie* cast doubt as to whether *pa* attaches to the entire constituent or not. Crucially, following Mathieu & Sitaridou (2005) analysis on split-DPs, the splitting is ultimately the result of movement which is driven by discourse considerations namely emphasis/contrast. This is perfectly compatible with the contrastive reading of the split-*pa* phrase attested here and which will be advocated for in the next section.

Furthermore, *pa*-phrases in Pontic can be multiple⁶ albeit this is not very frequent:

⁵ According to Drettas (1997:448), the expression *aets-pa*, “this way” is a fixed expression –a calque in other words– which marks a rupture/discontinuity in terms of discourse.

⁶ Benincà (2004: 53, 71), however, claims that recursion in the CP domain is not an option. In light of her proposal, each projection bears a particular semantic property and can host only one XP. Moreover, the highest projections encapsulate old information while new information appears lower in the CP area which hosts three subfields: Frame, Thematisation, & Focus.

(19) Multiple *pa*-phrases
(Pontic)

a. *ekino-pa propayanða eton ekino-pa*
that-NOM Top-PART propaganda-NOM that-NOM Top-PART
'That was propaganda.'

(Drettas 1997: 586 ex.248)

b. *zante, e'inon-pa skotonen ke emas-pa*
fool, that-ACC Top-PART kill-3PL-Pres and us-ACC Top-PART
'You fool, they'll kill both him and us.'

(Tombaidis 1988: 91)

c. *esi-pa kala ke alos-pa ke alos-pa, eksegketen tin kutsin ke ula epiketen*
you-NOM Top-PART well and other-NOM Top-PART and other-NOM Top-PART,
come-out-3PL-Past the-ACC limp-person-ACC and everything take-3PL-Past

'You and the others pulled out the limp lady and took everything.'

(Tombaidis 1988:83)

In (19a), the subject topic, *ekino-pa* "that" is recursive emerging both clause-initially and clause-finally thus enhancing the emphatic reading of the clause. On the other hand, (19b) illustrates an instance of multiple topics as the result of coordination: the two coordinated object DPs, preverbal *e'inon-pa* "that" and post-verbal *emas-pa* "us". Likewise, in (19c) all *pa*-topics are subject DPs, each with its own *pa*-marker and are all merged preverbally. However, the most genuine instance of multiple *pa*-topics is the one in (20) whereby multiple *pa*-constituents have distinct syntactic functions within the same clause:

(20) *kj atot eraepsanaton; ekin-pa ekints-pa efayane* (Pontic)

and then seek-3PL-Past+Obj-3SG-Masc this-Deict-NOM Top-PART this-Deict-ACC
Top-PART eat-3PL-Past

'And then they kept searching for those who killed the others.'

(Drettas 1997:440 ex.102)

Both the subject, *ekin-pa* "that" and the object, *ekints-pa* "those" of the second main clause bear the *pa*-marker and occur preverbally.

5. Syntactic analysis of topicalisation in Pontic Greek

In the literature there are different proposals regarding the division of labour between the distinct components of the grammar which are involved in the organisation of information structure. With regards to the actual mapping between the syntax and the interfaces there are *grosso modo* two main approaches: the feature-driven one and the stress-based one. Here we endorse the former. The feature-based approaches suggest a direct and unambiguous mapping between the grammatical representation of an utterance and its discourse interpretation. According to the Mapping Hypothesis, as developed by Diesing (1992), there is a strict correspondence between the syntactic form and the semantic interpretation. Within the feature-based proposals, it is the Cartographic Project (Rizzi 1997, 2004, Cinque 1999, 2002, 2006, Belletti 2004), and, in particular, the split-CP representation of the left-periphery of the sentences (cf. Rizzi 1997), which we adopt here. Although, the extension of the CP domain with additional functional heads has not been welcomed by more restrictively defined minimalist accounts (cf. Cormack & Smith 2000), our choice of framework is guided by the nature of the data.

Let us start by considering how high up on the syntactic tree *pa*-topics are located. First, consider (21):

(21) *I petheram-pa sin eklisian efeven* (Pontic)

The-NOM mother-in-law-NOM+Poss Top-PART to-the-ACC church-ACC go-3SG-Past
My mother-in-law went to the church

(Andreadis 1990:20)

(21) shows that *pa* is merged after the enclitic possessive *-m* clearly excluding any possibilities of *pa* being merged within the DP.

Second, consider diagnostics involving adverbial placement (cf. Cinque 1999) (22):

(22) a. “always”

T’emon o Jagon-pa panda estil’ne (Pontic)
 the mine the John-NOM Top-PART always sent
 ‘My boy, John was always sending stuff’
 (Andreadis 1990:28)

b. “maybe”

tæmæk ato pa ekripsen
 Maybe this-ACC Top-PART hide-past-3SG
 ‘Perhaps he also hid this’
 (Drettas 1997:569 ex. 185)

Interestingly, (22a) shows that the *pa*-phrase is higher than *Asp_{Perfect}* whereas in (22b) it is below the speaker oriented adverb *tæmæk* “maybe”.

Third, (23) also suggests a high position of the *pa*-phrase because it appears higher than the mood particle *na* which is hosted in the lower CP domain.

(23) Na yazanev’s pola paradas ke ti manas-pa na min anaspalts
 (Pontic)

Mod-PART gather-2SG many money and the mother-ACC
 Top-PART Mod-PART neg forget-2SG
 ‘To make a lot of money and not to forget your mother’
 (Andreadis 1990:22)

The topicalised object *ti-manas-pa* “the mother” is merged higher than the modality particle *na* which, according to Roussou (2000), is in Cop.

Fourth, (24) dissolves any uncertainty with regards to the high position of *pa*-phrases:

(24) ato-pa pos erθen so nu-s’? (Pontic)
 that-NOM Top-PART how come-3SG-Past to-the-ACC mind-ACC your
 ‘How did that cross to your mind?’
 (Andreadis 1990: 33)

(24) shows that *atos-pa* “he” is merged above the *wh*-constituent. This clearly indicates how “high” up in the tree *pa*-phrases are located in Pontic.

However, although we have demonstrated that *pa*-phrases are in the CP the question as to whether they pertain to a specialised projection, such as ContrastiveTopicP, or not remains open. In fact, there are several problems with such a claim. First, consider an example of postverbal topicalisation (25):

(25) lejme pios ef paraðas? (Pontic)
 say-3SG-Pres+Obj-1SG-ACC who-NOM have-3SG-Pres money-ACC?
 Leo k eyo pa ...ekinos ekserætsen ults ekser kata onoman ke kata jenean
 say-1SG-Pres and-EMPH I-NOM Top-PART...he-Deictic-NOM know-3SG-
 Pres+Obj-3PL-ACC everyone-ACC know-3SG-Pres by name-ACC and by generation-ACC
 ‘S/he says: Who has money?
 And I say: Everyone knows him by name and generation.’

In (25) it should be clarified that the *k* “and” is not a coordinator but has an emphatic use, and, thus, it enhances a marked nature at the clause.

Second, consider multiple *pa*-phrases (26):

(26) *kj atot eraepsanaton; ekin-pa ekints-pa efayane* (Pontic)
 and then seek-3PL-Past+Obj-3SG-Masc this-Deict-NOM Top-PART this-Deict-ACC
 Top-PART eat-3PL-Past
 ‘And then they kept searching for those who killed the others.’
 (Drettas 1997:440 ex.102)

(26) is an instance of multiple *pa*-phrases which are perfectly possible in Pontic and this should cast doubt on a single, dedicated projection.

Finally, consider (27) which shows attachment of *pa* to a QP:

(27) *ul’ i kaloer’-pa ayapune ton Yorika-m’* (Pontic)
 all the-NOM monk-PL-NOM Top-PART love-3PL-Pres the-ACC George-ACC my
 ‘All the monks love my George.’
 (Melanofrydis 2001:29)

In (27) *pa* is attached to the right of the quantifier which is trivially assumed to be in the Spec-CP. If *pa* was a contrastive topic marker why is it possible to select a QP –an element which is inherently focalised? Overall, examples (25-27) provide counterarguments for a dedicated ContrastiveP projection.

On the basis of the arguments presented so far, we think there is enough evidence to dismiss the possibility for a dedicated ContrastiveTopic projection. Crucially though, this does not amount to claiming that we dismiss the idea of a specially designated position for *pa*-phrases or that *pa*-phrases cannot function as contrastive elements. Indeed, we claim that topicality and the contrastive interpretation associated with *pa*-phrases are two independent features of a contrastive topic, and thus agreeing with Vermeulen (2008). In what follows we present evidence for corroborating such a claim.

Let us start by examining the topicalisation strategy in two different languages namely, Japanese and Pontic Greek, which, however, both employ particles. Consider Table 2:

Table 16: Comparison of Japanese *wa*-phrases and Pontic *pa*-phrases

Properties	Japanese <i>wa</i> -phrases	Pontic <i>pa</i> -phrases
Multiple topics	Yes (but only one contrastive)	Yes
Particle as the only way of marking topics	Yes	No
Dedicated ContrastiveTopicP?	No	Possibly
Restrictions as to which category the particle attaches?	No (but not with predicates)	No (but not with predicates)
Focus markers also available	No	Yes

In Japanese the existence of a specialised particle such as *wa* has been taken as strong evidence for the existence of a Topic projection, *wa* being a morphological realisation of the Topic head. However, recent works (cf. Benincà and Poletto 2004, Frascarelli and Hinterhölzl 2007) also argue in favour of a syntactic encoding of different topic categories, and, in particular, they postulate a dedicated projection in the left periphery of the sentence for each type of topic. More specifically, Frascarelli and Hinterhölzl (2007) reject Rizzi's (1997) recursive definition of the Topic Phrase and propose the following topic hierarchy instead:

(28) Topic hierarchy

Shifting topic [+aboutness] > Contrastive topic > Familiar topic

In this hierarchy, three distinct projections are identified and, according to Frascarelli and Hinterhölzl (2007), each projection is associated with specific structural properties as well as different tonal events. Crucially, in Japanese, sentences containing multiple *wa*-phrases (29) are possible and sound most natural if there is no more than one non-contrastive *wa*-phrase, but there can be multiple contrastive *wa*-phrases (Kuno 1973, see also Kuroda 1988, Tomioka 2007 in Vermeulen 2008:20).

(29) a. sono inu-wa BILL-WA moo sudeni kyonen kandeiru. (Japanese)

that dog-wa Bill-wa already last.year bite-perf.

b. BILLi-WA sono inu-wa moo sudeni kyonen ti kandeiru.

Bill-wa that dog-wa already last.year bite-perf.

'That dog has already bitten Bill last year.'

(Vermeulen 2008: 1)

In general, the existence of contrastive topics has, as anticipated, important repercussions on the realisation of contrastive foci since in the literature contrastive topics are sometimes called foci, despite their thematic nature, thus, contributing to the blurring between the notions of focus and topic. In Finnish, for instance, contrastive focus and contrastive topic occupy the same structural designated position (cf. Vilkkuna 1995). Could this be also the case in Pontic? To put differently, is it possible that *pa* -to which we have referred as "contrastive topic marker"– is not a topic marker but rather a contrastive marker which can also function as a topic? Let us start by employing Rizzi's (1997) diagnostics between topics and foci in order to establish whether *pa*-phrases are topic-like or focus-like. First, consider the compatibility of *pa*-phrases with a resumptive clitic (30):

(30) Resumptive clitic

(Pontic)

Ato-pa pos epikesato?

this Top-PART how do-2SG-PAST-it

'How did you do this (and not something else)?'

(Andreadis 1990:54)

Second, consider *pa*-phrases which give rise to Weak Cross-Over without resulting to ungrammaticality:

(31) WCO

(Pontic)

Ton Jorikan-pa i manat pola ayap'aton

The George-ACC Top-PART the mother-NOM much love-3SG-him

'His mother loves George a lot'

Third, consider bare quantificational elements which can take *pa*-marking (32):

(32) Bare quantificational elements

(Pontic)

Ke ul-pa ekatetyoresan'aton
 and all-NOM Top-PART blame-3PL-PAST-him
 'And they all blamed him'
 (Melanofrydis 2001:41)

Fourth, multiple *pa*-phrases are possible, as we have already seen in (26) –repeated here as (33) for convenience:

(33) Multiple *pa*-phrases
 (Pontic)
 kj atot eraepsanaton; ekin-pa ekints-pa efayane
 and then seek-3PL-Past+Obj-3SG-Masc this-Deict-NOM Top-PART this-Deict-ACC Top-PART eat-3PL-Past
 'And then they kept searching for those who killed the others.'
 (Drettas 1997:440 ex.102)

Finally, consider the compatibility of *pa*-phrases with *wh*, as in (34):

(34) Compatibility with *wh* (Pontic)
 T'atines-pa ta tærta pjos apori na sirata
 Her own chagrins who can tolerate
 (Andreadis 1990:17)

To summarise our findings so far, consider Table 3:

Table 17: Comparing Pontic *pa*-phrases to topics and foci

Properties	Focus	<i>pa</i> -phrase	Topic
Resumptive clitic	No	Yes	Yes
Weak Cross-Over	No	Yes	Yes
Bare quantificational elements	Yes	Yes	No
Uniqueness	Yes	No (but only marginally so)	No
Compatibility with Wh	No	Yes	Yes

Table 3 clearly shows that *pa*-phrases show a mixed behavior: sometimes behaving like foci and others like topics. In order to resolve this odd behavior we must briefly consider the articulation of focus, and, in particular, contrastive focus for which discourse particles are also used (Kaltsa & Sitaridou, to appear). Consider (35):

(35) ar aets pontiaka peaton-ki na esker (Pontic)
 so that-way pontic-ACC say+Obj-3SG-ACC Foc-PART Mod-PART know-3SG-Pres
 'Hence, tell him in Pontic so that he knows.'
 (Drettas 1997: 523 ex.5)

In (35) the *ki* particle appears attached after the fused verb/object and focalises the entire TP. Kaltsa & Sitaridou (to appear) claim that *ki* does not attach enclitically to any

other element except for predicates, namely Heads. For this reason, the following example would be ungrammatical:

- (36) *Tin Anasta-ki iða (not Partheni) (Pontic)
 The-ACC Anasta-ACC Foc-PART see-Past-1SG
 'I saw ANASTASIA (not Parthena)'

Crucially, we have seen in section 4 that *pa* attaches to XPs but not to X⁰s. We would therefore, like to suggest that *pa* and *-ki/kela* are in complementary distribution and that they are therefore, realised in a single projection which we will call ContrastiveP. To put it differently, what we have called so far contrastive focus particles and contrastive topic particles we would like to propose that they are merely contrastive. In our analysis topicality and contrastivity are two independent features.

Let us now consider the proposed articulation of information structure in Pontic, as it emerges from the discussion so far. The empirical generalisation is that Pontic has *regressione del nuovo* "regression of the new" contrary to Greek and most of the Romance languages which have *progressione del nuovo* "progression of the new" (Benincà and Salvi 1988: 118-119). Consider the orderings in (37) which give an insight to the overall articulation of the information structure in Pontic:

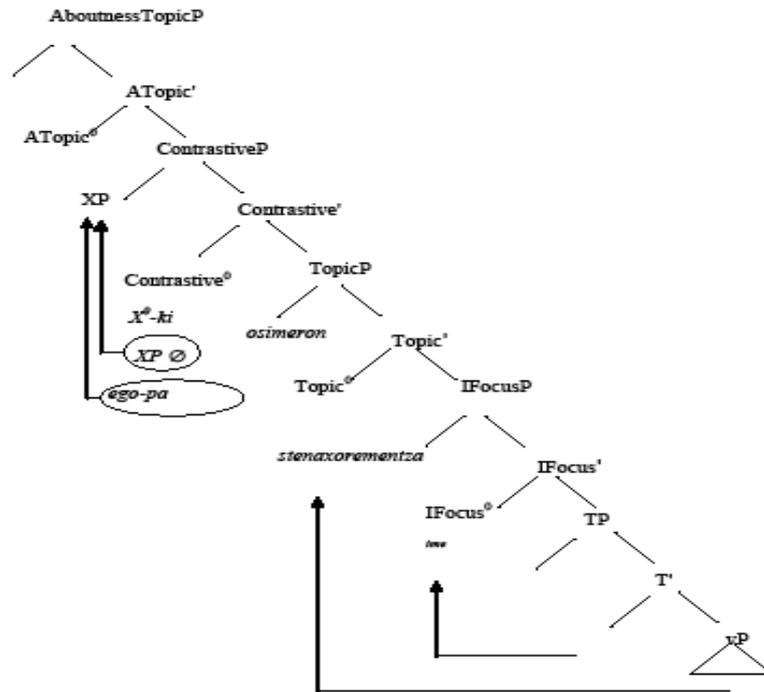
- (37) a. CLLD Object – Subject-*pa* – V (Pontic)
 Ton Memet ego-*pa* agapoaton
 the-ACC Memet-ACC Pronoun-1SG+Particle love-1SG-Pres+Pronoun-ACC
 It is Memet that I love
 (Melanofridis 2001:13)
- b. Subject – Object-*pa* – IFoc – V
 i Nazlu-xanum ekinon-*pa* efkero ki θ' afin'
 the-NOM Nazlu-xanum-NOM that-Deict-ACC Top-PART empty-ACC Neg-PART
 Fut-PART leave-3SG
 'Nazlu-xanum wouldn't leave that empty.'
 (Melanofrydis 2001: 43)
- c. Object-*pa* – IFoc – V
 k ekina-*pa* o popas eton
 and those-PART the-NOM priest-NOM be-3SG-Past
 (Drettas 1997:442)
- d. Subject-*pa* – Topic – IFoc – V
 Ego-*pa* osimeron pola stenaxorementza ime
 I-PART today many sad be-1SG-Pres
 'Today I am so sad.'
 (Andreadis 1990:27)

The above examples suggest the following hierarchy:

- (38) (Aboutness)TopicP ... ContrastiveP (Topic) ... IFoc ... TP

The structure we assume to be at work is shown in (39):

- (39)



(39) suggests that in the Head of ContrastiveP there can be merging of any one of three different lexical heads: *-pa*, *zero*, *-ki*. *Pa* and \emptyset select XPs, and therefore, they move to the specifier of ContrastiveP. *Ki*, on the other hand, by virtue of selecting X^0 does not move further. *Pa*-phrases by virtue of being in the Spec-ContrastiveP may be interpreted as topics. Additionally, there can be Topic projections above ContrastiveP and between the latter and IFoc.

To conclude, consider the parametric variation between Pontic and SMG, as shown in Table 4:

Table 18: Information structure and (micro)parametric variation in Greek

Pontic	SMG
Specific structural positions for topics, contrastive elements and information focus in the left periphery	Specific structural positions for topics and contrastive focus in the left periphery whereas information focus is in the right one
Morphological encoding of discourse	No morphological encoding of discourse

6. Conclusion

In this paper we have argued that the information structure of Pontic Greek is organised in a radically different way from the one in SMG by virtue of making extensive use of particles. More concretely, we argued in favour of a contrastive projection in the CP domain which can host both topics and foci. *Pa* is argued to select XPs which can then be interpreted as topics hence why all *pa*-constituents would receive the reading of “contrastive topics”.

References

Adreadis, R. (1990). *Pontiako Theatro: I Ksenitia*. Thessaloniki.
 Belletti, A. (2004). Aspects of the low IP area. In Luigi Rizzi (ed.) *The Structure of IP and CP. The Cartography of Syntactic Structures Vol. 2*: 16-51. Oxford/New York: OUP.

- Benincà, P. & C. Poletto (2004). Topic, focus and V2: defining the CP sublayers. In Luigi Rizzi (ed.) *The Structure of IP and CP. The Cartography of Syntactic Structures, Vol. 2*: 52-75. Oxford/New York: OUP.
- Benincà; Paola and Giampaolo Salvi (1988). L'ordine degli elementi della frase e le costruzioni marcate (§1.1). In L. Renzi & G. Salvi (eds.) *Grande grammatica italiana di consultazione Vol. 1*: 119-129. Bologna: Il Mulino.
- Brody, M. (1990). Remarks on the order of elements in the Hungarian focus field. In I. Kenesei (ed.) *Approaches to Hungarian: structures and arguments 3*: 207-243 JATE: Szeged..
- Chomsky, N. (1976). Conditions on rules of grammar. *Linguistic Analysis 2*: 303-350.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Cinque, G. (2006). (ed.) *Restructuring and Functional Heads. The Cartography of Syntactic Structures, Vol. 4*. Oxford/New York: OUP.
- Cinque, G. (2002). *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Vol. 1*. Oxford/New York: OUP.
- Cinque, G. (1999). *Adverbs and Functional Heads*. Oxford/New York: OUP.
- Cinque, G. (1990). *Types of A'-dependencies*. Cambridge, Mass: MIT Press.
- Cormack A. & N. Smith (2000). Fronting: the syntax and pragmatics of 'focus' and 'topic'. *UCL Working Papers 12*: 387-416.
- Culicover, P. (1991). Topicalisation, inversion, and complementizers in English. In D. Delfitto et al. (eds.) *Going Romance and beyond: fifth symposium on com-parative grammar*: 1-43 Utrecht: Utrecht University.
- Dawkins, R. M. (1916). *Modern Greek in Asia Minor: A study of the dialects of Silli, Cappadocia and Pharsa with grammar, texts, translations and glossary*. Cambridge: CUP.
- Dezsö, L. (1974). Topics in syntactic typology. In M. Romportl. et al. (eds.) *Linguistica Generalia: Studies in Linguistic Typology*, 191-210. Prague: Charles University.
- Diesing, M. (1992). *Indefinites*. Cambridge, Mass: MIT Press.
- Drettas, G. (1997). *Aspects Pontiques*. Association de recherches pluridisciplinaires.
- Drettas, G. (2000). Marques de focus en grec commun et en pontique, *Studies in Greek Linguistics*, Département de Linguistique de l'Université de Salonique, 122-131
- Erteschik-Shir, N. (1997). *The Dynamics of Focus Structure*. Cambridge: CUP.
- Erteschik-Shir, N. & N. Strahov (2004). Focus structure architecture and P syntax. *Lingua 114* (3): 301-323.
- Frascarelli, M. & R. Hinterhölzl (2007). Types of Topics in German and Italian. In S. Winkler & K. Schwabe (eds.) *On Information Structure, Meaning and Form*: 87-116. Amsterdam: John Benjamins.
- Hajičova, E., Patree, B. & P. Sgall (1998). *Topic-focus articulation: tripartite structures and semantic content*. London: Kluwer.
- Hajičova, E. & P. Sgall (1988). Topic and focus of a sentence and the patterning of a text. In J. Petöfi (ed.) *Text and discourse constitution*: 70-96. Berlin: De Gruyter.
- Horvath, J. (1986). *Focus in the Theory of Grammar and the Structure of Hungarian*. Dordrecht: Foris.
- Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.
- Huang, J. (1984). Move WH in a language without WH movement. *Linguistic Review 1*: 369-416.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, Mass: MIT Press.
- Janse, M. (2002). Aspects of Pontic grammar. *Journal of Greek linguistics 3*: 203-231.
- Kaltsa, M (2007). *Information Structure in Varieties of Greek*. Unpublished MPhil Thesis, University of Cambridge, UK.
- Kaltsa, M & I.Sitaridou (to appear). *The organisation of information structure in Pontic*. Ms. available, AUTH & University of Cambridge.
- Kiss, K. E. (1998). Identificational focus versus informational focus. *Language 74*/2: 245-273.
- Koizumi, M. (1994). Layered specifiers. *NELS 25*: 255-269.
- Koizumi, M. (1995). *Phrase structure in minimalist syntax*. Unpublished PhD. Thesis, MIT University.
- Lambrecht, K. (1994). *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge: CUP.
- Mackridge, P. (1996). 'The Medieval Greek Infinitive in the Light of Modern Dialectal Evidence' in C. N. Constantinides et al., ΦΙΛΕΛΛΗΝ. Studies in honour of Robert Browning. Venice: Istituto ellenico di studi bizantini e postbizantini di Venezia, 191-204.
- Mackridge, P. (1995). Τα Ποντιακά στη σημερινή Τουρκία: αρχαία στοιχεία στο ιδίωμα του Όφης. *Αρχαίον Πόντου 46*: 153-161.
- Michelioudakis, D. & I. Sitaridou (in prep.). Double object constructions in Pontic. Ms. available,

- University of Cambridge.
- Melanofrydis, P. (2001). *I klosti*. Thessaloniki: Adelfon Kiryakide.
- Neeleman, A. & T. Reinhart (1998). Scrambling and the PF-interface. In M. Butt & W. Geuder (eds.) *The projection of arguments: lexical and compositional factors*: 309-353. Chicago: CSLI Publications.
- Papadopoulos, A. (1955). *Istoriki grammatiki tis pontikis dialektu*. Athens: Epitropi Pontiakon Meleton.
- Papadopoulos, A. (1958-61). *Istorikon lexicon tis pontikis dialektu*. Athens: Epitropi Pontiakon Meleton.
- Pappas, P. (2006). 'Object Clitic Placement in the Dialects of Medieval Greek'. In M. Janse, B. Joseph & A. Ralli (eds.) *Proceedings of the Second International Conference of Modern Greek Dialects and Linguistic Theory*: 314-328. Mytilene: University of Patras Press
- Ralli, A. (2007). Syntactic and morpho-syntactic phenomena in Modern Greek dialects. *Journal of Greek linguistics*.
- Reinhart, T. (1995). *Interface strategies*. Utrecht: OTS.
- Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* **27**: 53-94.
- Revithiadou, A. (2008). "A cross-dialectal study of cliticization in Greek." In C. Ewen, H. Van der Hulst & N. Kula (eds.) *Lingua. Special Issue in memory of J.G. Kooij*,
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (ed.) *Elements of grammar: handbook in generative syntax*: 281-337. Dordrecht: Kluwer.
- Rochemont, M. (1989). Topic Islands and the subjacency parameter. *Canadian journal of linguistics* **34**: 145-170.
- Rousou, A. (2000). On the Left Periphery: Modal Particles and Complementisers. *Journal of Greek Linguistics* **1**: 65-94.
- Setatos, M. (1994). «Το ποντιακό πα». *Ελληνική Διαλεκτολογία* **4**. Θεσσαλονίκη: Αφοί Κυριακίδη.
- Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic inquiry* **31.4**: 649-689.
- Szendrői, K. (2004). Focus and the interpretation between syntax and pragmatics. *Lingua* **114**: 229-254.
- Tombaidis, D. (1988). *I pontiaki dialektos*. Athens: Epitropi Pontikon Meleton.
- Tombaidis, D. E. (1998). Πάλι για το απαρέμφοτο στην Ποντιακή διάλεκτο. *Αρχεῖον Πόντου* **48**: 5-14.
- Vallduví, E. (1993). Informational packaging: a survey. MS., University of Edinburgh.
- Vermeulen, R. (2008). Topics in Japanese: A unified analysis of contrastive and non-contrastive topics. Ms. UCL
- Vermeulen, R. (2007). Non-topical *wa*-phrases in Japanese. A revised paper version of the talk presented at OnLI conference in Belfast, June 2007. Ms. UCL
- Vilkuna, M. (1995). Discourse configurationality in Finnish. In K. É. Kiss (ed.) *Discourse-Configurational Languages*: 244-268. Oxford/New York: OUP.
- Winkel, S. & E. Göbbel (2002). Focus, p-movement, and the nuclear-stress rule: a view from Germanic and Romance. *Linguistics* **40.6**: 1185-1242.
- Zubizarreta, M. L. (1998). *Prosody, focus and word order*. Cambridge, Mass: MIT Press.

Clitic climbing in Grecia Salentina Greek: A dynamic account

STERGIOS CHATZIKYRIAKIDIS

King's College, London

kafouroutsos@hotmail.com

1. Introduction

Grecia Salentina Greek (GSG) is one of the two varieties of Greek spoken in Southern Italy, the other one being Calabrian Greek. GSG is spoken in a number of villages¹ around the area of Lecce, Southeast Italy.

The clitic system of GSG resembles to a high degree that of Standard Modern Greek in terms of positioning. Clitics are proclitic with non-imperative verbs and enclitic with imperatives:

- (1) *Ton gapa*
him.CL-ACC loves
'He loves him.'
- (2) **Gapa ton*
loves him.CL-ACC
'He loves him.'
- (3) *Grafe to*
write.IMP-2SG it.CL-ACC
'Write it!'
- (4) **To grafe*
it.CL-ACC write.IMP-2SG
'Write it!'

The only difference noted in the literature (Ralli, 2006) is that GSG unlike SMG does not allow free ordering in clitic clusters in imperative environments:

- (5) *Δos mu to*
give.IMP-2SG me.CL-GEN it.CL-ACC
'Give it to me.' [SMG]
- (6) *Δos to mu*
give.IMP-2SG it.CL-ACC me.CL-GEN
'Give it to me.' [SMG]

¹The exact number is hard to define. The standard assumption is that the number is 9. However, in a number of these villages like e.g. Zollino GSG is not spoken anymore.

(7) *Do mu to*
 give.IMP-2SG me.CL-GEN it.CL-ACC
 ‘Give it to me.’ [GSG]

(8) **Do to mu*
 give.IMP-2SG it.CL-ACC me.CL-GEN
 ‘Give it to me.’

GSG is also the only Modern Greek dialect exhibiting clitic climbing with non-auxiliary verbs, a fact that has not been noticed in the literature so far. In specific, GSG allows clitic climbing with two verbs of the restructuring class,² the verbs *sotzo* and *spitseo*, ‘can’ and ‘finish’ respectively. However, unlike CC in languages like Italian (Cinque 2002, 2006; Cardinaletti & Shlonsky, 2004 among others), CC is obligatory in GSG:

(9)

a. *To sotzume avorasi.*
 it.CL-ACC can buy.INF

b. **Sotzume to avorasi.*
 can it.CL-ACC buy.INF

c. **Sotzume avorasi to.*
 can buy.INF it.CL-ACC
 ‘We can buy it.’

(10)

a. *To spitseo tse torisi avri*
 it.CL-ACC finish.1SG COMP see.INF tomorrow

b. **Spitseo tse to torisi avri*
 finish.1SG COMP it.CL-ACC see.INF tomorrow

c. **Spitseo tse torisi to avri*
 finish.1SG COMP see.INF it.CL-ACC tomorrow
 ‘I will finish seeing it tomorrow.’

It is crucial to note that these two verbs are the only ones of the restructuring class that still subcategorize for an infinitive. Such a fact is crucial in understanding the unavailability of climbing with the rest of the verbs of the same class, since all the other restructuring verbs use the subjunctive marker *na* as the complementation strategy:

(11)

a. *Telume no(na-to) avorasume.*
 want SUBJ-it.CL-ACC buy.1PL

b. **To telume na avorasume.*
 it.CL-ACCwant SUBJ buy.1PL
 ‘We want to buy it.’

²Restructuring verbs include modal, motion and aspectual verbs.

(12)

- a. *Ancigneo na to toro*
 begin.1SG SUBJ it.CL-ACC see.INF
- b. **To Ancigneo na toro*
 it.CL-ACC begin.1SG SUBJ see.INF
 ‘I begin to see it.’

It seems that the unavailability of the infinitival strategy in the rest of the verbs of the restructuring class is at least one of the reasons that climbing is only available for just these two verbs. Assuming monoclausality is what really lies behind CC constructions, the subjunctive strategy is incompatible with a monoclausal interpretation since disjoint reference is always possible given that the verb following the subjunctive marker is finite. What other semantic or historical reasons may be behind this rather idiosyncratic property of GSG is something that I do not know. However, it does not seem that anything semantic is really at play here. There is no generalization across classes and verbs similar in meaning (compare *spicceo* and *ancigneo* for example) behave differently with respect to climbing. Surprisingly, Romanian which like GSG uses the subjunctive strategy but retains the infinitive with a limited number of verbs, does not allow climbing except with the verb *a putea* ‘can’:

(13)

- a. *O pot vedea*
 her.CL-ACC can see.INF
- b. **Pot vedea o*
 can see.INF her.CL-ACC
 ‘I can see her.’

I do not have any historical story of why some verbs retained the infinitive and some did not, and as such this issue will not be discussed here. On the other hand, I will try to provide an account of obligatoriness of CC in GSG within the DS framework. But before doing that, let us first take a look at the most prominent analyses of CC in the literature.

2. Approaching Clitic Climbing

The first thing one should think of when giving an analysis of CC, is what is the problem with such constructions, i.e. what makes them problematic to linguistic theory. The problem can be simply stated as a locality violation, with clitics being attached to a verbal host other than the one they constitute arguments of. A number of different accounts have been proposed by the years in different grammatical frameworks. Earlier approaches within the GB/Minimalist tradition assume that CC is a case of restructuring. In all these approaches, a restructuring rule is invoked, in effect transforming a biclausal structure into a monoclausal one. Such an approach can be found for example in Rizzi (1983), where a restructuring rule is posited to account for CC (see also Manzini (1983) for a similar treatment):

(14) Rizzi’s restructuring rule - My formalization

$$V (P) V.INF \rightarrow V.CMLX$$

The above rule transforms a biclausal structure consisting of a verb and an infinitive into a monoclausal structure where the two verbs are assumed to form a verbal complex. However, the fact that in some CC languages a number of adverbs can intervene between the verb and the infinitive caused serious problems to such accounts. On the other hand, Kayne (1989) working within the GB/Barriers framework claimed that

CC is the result of clitic movement out of the infinitival VP. The exact reasoning is then that in non-CC languages the VP constitutes a barrier for movement and thus movement is debarred, while in CC languages the IP L-marks the VP and the latter is not considered a blocking category anymore (in the sense of Chomsky, 1986).

Recent approaches within the minimalist program propose that clitic climbing occurs when one of the verbs appears not as a fully fledged verb heading its own VP, but rather as an instantiation of an FP within the richly articulated FP structure of the clause proposed by Cinque (1999). In that sense optional climbing is caused when one verb can be inserted either as a functional or a lexical verb:

(15) Functional and lexical instantiation of a verb

- a) [CP...[FP...[FP V_{restr} [FP...[VP V]]]] Climbing case
- b) [CP...[FP...[FP[VP V_{restr} [CP...[FP...[FP[VP V]]]]]] Non-climbing case

Cardinaletti and Shlonsky (2004) propose that a) is involved in CC constructions and b) in non-CC constructions. Cinque (2006) on the other hand argues that a) is involved in both cases.

Within HPSG, CC has been considered to be an argument sharing phenomenon (Miller and Sag 1997, Monachesi (1998, 1999) among others). All the HPSG analyses concur on the latter claim. The assumption is that the climbing inducing verb subcategorizes for an infinitive plus its arguments:

$$(16) \left[\begin{array}{l} \text{HEAD } V \\ \text{VCLASS modal } \vee \text{ aspectual } \vee \text{ motion} \\ \text{SUBJ} \langle NP \rangle \\ \text{COMPS } L \oplus \quad \langle V \left[\begin{array}{l} \text{CLTS } \{ \} \\ \langle NP \rangle \\ \text{COMPS } L \end{array} \right] \end{array} \right]$$

Argument sharing explains why the clitic can climb in CC constructions but does not however have anything to say with respect to restructuring effects found in CC environments like for example unavailability of infinitival negation when CC has taken place (example below from Italian):

(17) **Lo vuole non vedere*
 it.CL-ACC want NEG buy.INF
 'I want to not see it.'

(18) *Vuole non vederlo*
 want NEG buy.INF it.CL-ACC
 'I want to not see it.'

Furthermore, it is not clear to me what subcategorization for an infinitive plus its arguments means and furthermore why non-restructuring verbs are not able to do this kind of subcategorization.

I will not go into the exact details of all these approaches since it is not my attention to refine on any of these analyses, but rather to provide an alternative DS analysis. Therefore, in what follows I will present a DS analysis of CC, arguing that the functional vs non-functional distinction assumed in the recent minimalist literature can receive a better interpretation and formalization once a shift towards a parsing oriented framework is done. But first an introduction to the Dynamic Syntax (DS) framework.

3. An informal introduction to Dynamic Syntax

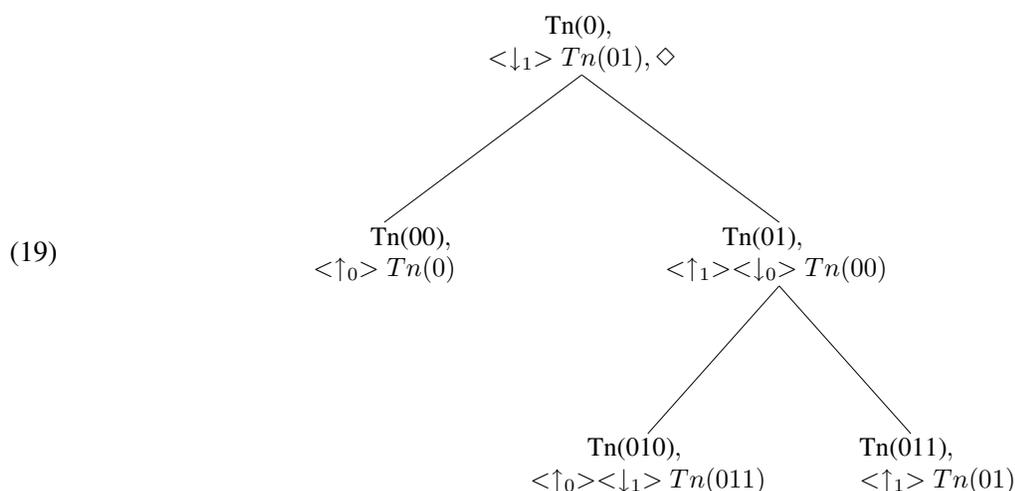
3.1 Basic intuitions behind Dynamic Syntax

The basic assumption behind DS is that natural languages are interpreted via an incremental, word-to-word, left-to-right cumulative construction of transparent semantic representations with the upper goal to finally construct a logical form of type t ($?Ty(t)$). Such an interpretation is driven by means of monotonic tree growth, representing the attempt to model the way information is processed in a time-linear, word-to-word manner. However, tree-structures in DS are considerably different from those found in derivational or declarative frameworks like minimalism or HPSG respectively, in that they are not inhabited by words as such, but rather from the representations of those words (Fodor, 1975). Furthermore, the tree structure corresponding to the representation of the ending result of parsing a natural language string is a semantic representation assigned to this natural language string with respect to some context. This semantic representation does not correspond to word order but rather represents argument structure. However, the incremental left-to-right parsing via an array of successive, monotonically growing tree structures, handles word order through the mere definition of incremental parsing. The partial tree structures or the history of parsing stages are used to capture word order phenomena, since this whole process is totally dependent on the way words are ordered. In order for all these intuitions to be carried out, a number of formal tools are employed.

3.2 The formal framework in a glance

3.2.1 LOFT, Tree decorations, Requirements

The parsing process is represented by means of binary trees underpinned by the Logic Of Finite Trees (LOFT, Blackburn and Meyer - Viol, 2001). Left branches are addressed conventionally by adding 0 to the value of the mother node, while right branches by adding 1. The position of a given node is expressed using the predicate Tn (standing for treenode) followed by the actual treenode address. Furthermore two basic tree modalities, $\langle \uparrow \rangle$ and $\langle \downarrow \rangle$, standing for the mother and daughter relationship respectively, allow relations between the trees to be represented:



Notice that a given treenode can be addressed from the perspective of a different treenode. For example $\langle \uparrow_0 \rangle \langle \downarrow_1 \rangle Tn(011)$ in the 010 node reads as follows: You will find treenode 011 if you first go up the 0 mother relation and then go down the 1 daughter relation. The \diamond symbol, found in the 0 node in our example is called the pointer, and its basic function is to track which node is the current node under construction any time during the parsing process.³ Nodes in DS are inhabited by a set of labels, conventionally called “Tree Decorations”. The basic elements comprising this set are:

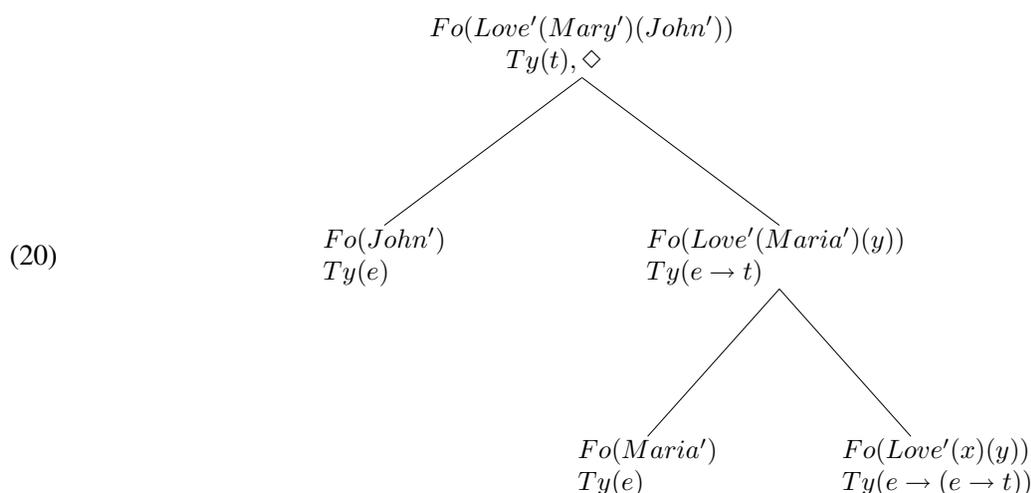
- a. **Formula value decorations.** These are represented using the predicate Fo followed by the representation of the entity in brackets, e.g. $Fo(\text{John}')$.⁴

³The Pointer function is also one of the ways to account for ungrammaticality in DS. For instance, if the pointer is at a given node which has an outstanding requirement for a type e expression to be found, and the next word that is parsed does not provide such a Type, by providing e.g. a $Ty(e \rightarrow t)$ expression, the parse will abort rendering the whole string ungrammatical.

⁴The prime indicates that the concept and not the word John is represented.

- b. **Type value decorations.** These are represented using the predicate *Ty* followed by the type of the word/concept in question in brackets, e.g. *Ty*(e).⁵

A basic concept in the DS framework is that of requirements. Requirements can be seen as goals that need to be achieved. Requirements have the general form $?La_i$ ⁶ (e.g. $?Ty(e)$). In order for a given parse to be successful, no outstanding requirements must exist in the ending tree. In that respect, requirements can be also seen as a device explaining ungrammaticality. Example (19) shows a complete tree in DS. Notice that no outstanding requirements exist⁷:



3.2.2 Computational - Lexical - Pragmatic rules/actions

The parsing process is driven by three kinds of rules/actions: a) Computational b) Lexical and c) Pragmatic rules/actions. The former are general computational devices, comprising the basic tree construction mechanism. They always involve an input and an output description. The former designates where the pointer must be along with information about the node that the pointer is on or other nodes with respect to the pointer node, while the latter shows the transformation of the input in terms of requirements, adding nodes, pointer movement etc. An example of a computational rule, the rule of COMPLETION, is shown below:

$$(21) \frac{\{...\{Tn(n),...\}, \{<\uparrow_i>, Tn(n),..., Ty(X) \dots, \diamond\}...\}}{\{...\{Tn(n), \dots, <\downarrow_i> Ty(X), \dots, \diamond\}, \{<\uparrow_i> Tn(n), \dots, Ty(X), \dots\}...\}} \quad 8$$

The above rule moves the pointer to the mother node as soon as any type of requirement is satisfied on any of the daughters of that mother node. The top part reads as follows: There is a node with treenode address $Tn(n)$ and another one dominated by $Tn(n)$ ($<\uparrow_i>, Tn(n)$) that has a satisfied type requirement and also bears the pointer. The output description (second line) presents a situation where the pointer has moved to the $Tn(n)$ addressed node, with an additional statement that records the daughter's satisfied requirement ($<\downarrow_i> Ty(X)$). There are a number of procedural computational rules like the one we have just seen but we won't go into the rest of them for reasons of space. The interested reader is referred to Kempson et al. (2001), Cann et al (2005) for a detailed presentation of a number of computational rules. Additional computational rules will be introduced if needed.

⁵The difference between DS and most of the formal semantic theories with respect to typing is twofold. Firstly, DS has an additional type (cn) standing for common noun, and furthermore there is no recursion on types. Types are a rather closed set. For a more detailed discussion on DS typing see Kempson et al. (2001) and Cann et al. (2005).

⁶Where La stands for label and $i \geq 1$. For a formal presentation of declarative units in DS consult Kempson et al. (2001: chapter 7).

⁷The lambda terms in the Fo formulas have been excluded for ease of exposition.

⁸Where $i=(0,1,*)$.

On the other hand, lexical rules are basically entries associated with a given word providing instructions on how the parsing must or must not proceed. There are no general rules regarding the content of these instructions. They rather depend on the syntactic nature of these words. However, there is a generalized schema involved in the way these words introduce their content. This general procedural schema followed by a sample lexical entry is shown below:

(22) Lexical entry format

IF Trigger
 THEN Actions
 ELSE Elsewhere statement

(23) Sample lexical entry for *Bill*

IF ?*Ty*(*e*)
 THEN put(*Ty*(*e*), *Fo*(*Bill'*), [↓]⊥)
 ELSE abort

Example (23) reads as follows: If you are in a node that has a type *e* requirement, then decorate this node with a type *e* value and a formula value representing the concept given by the word *Bill*. In any other case abort. A proper noun like *Bill* in English will be able to get parsed as soon as a node has a requirement for a type *e*. This will allow a word like *Bill* to be parsed either as a subject or as an object in English. Other languages with overt noun case marking will have further restrictions for their equivalent entry for *Bill* that will ensure that the proper noun will be parsed in the relevant node depending on case marking. For example we can associate a given case marking with a fixed structural position by means of tree modalities as shown below:

(24) Structural accusative lexical entry

IF ?*Ty*(*e*), ⟨↑₁⟩?*Ty*(*e* → *t*)
 THEN put(*Ty*(*e*), *Fo*(*x'*), [↓]⊥)
 ELSE abort

The above entry will block a noun of the above type to be parsed in the subject node (00) simply because the condition ?*Ty*(*e* → *t*) is not going to be satisfied.⁹

Lastly, pragmatic actions involve contextual information providing information with respect to the parsing process. One very basic rule is the rule of SUBSTITUTION which updates a formula metavariable into a proper formula referring to some entity in the context.¹⁰ I won't discuss any other pragmatic actions in this paper but the interested reader is directed to Kempson et al. (2001) and Cann et al. (2005) for further information on pragmatic actions.

3.2.3 Underspecification, LINK

Central within the DS framework is the concept of underspecification, the idea that parts of natural language may not be fully specified by the time they enter into the parsing procedure. Of course underspecification is not in itself a new concept in linguistics. It has been extensively used the last 20 years by formal semanticists to deal with ambiguity and anaphora resolution. What is novel however, is that underspecification is moved into the area of syntax,¹¹ making the syntax the dynamic part rather than the semantics. DS uses two types of underspecification: a) Content underspecification and b) Structural underspecification.

⁹In DS binary trees, as already mentioned, do not encode word order but rather represent argument structure. In that respect the subject node is always in the same position no matter what the actual word order is. This position is the 00 node. Given that, it is clear why the condition is not satisfied.

¹⁰See Kempson et al. (2001) and Cann et al. (2005) for a formal definition of the rule of SUBSTITUTION.

¹¹There is however a similar notion, the notion of functional uncertainty in LFG (Bresnan, 2001)

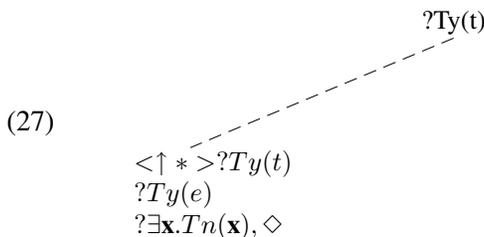
With respect to content underspecification, DS employs the use of metavariables, functioning as mere content placeholders with a requirement that substitution of the metavariable will take place at later stages of the parse. A classic example of content underspecification is pronouns. DS assumes that the lexical entry for a pronoun, say *he*, involves the projection of a metavariable as the value the Formula takes. This metavariable must be updated as soon as a proper formula value is provided by the context or by the natural language string itself. The metavariable comes with person and case restrictions depending on the pronoun. A requirement that a proper Fo value must be found ensures that the node which bears the metavariable will eventually get a proper formula value. The lexical entry for the pronoun *he* is shown below:

- (25) Lexical entry for *he*
- | | |
|------|--|
| IF | $?Ty(e), \langle \uparrow_1 \rangle ?Ty(t)$ |
| THEN | $put(Ty(e), Fo(U_{male'}), ?\exists \mathbf{x}. Fo(\mathbf{x}), [\downarrow] \perp)$ |
| ELSE | <i>abort</i> |

Structural underspecification on the other hand is represented in DS by employing a set of rules, the family of ADJUNCTION rules.¹² *ADJUNCTION effectively introduces a node, which position in the tree is not yet fixed at the time of its introduction. To be more specific, the rule of *ADJUNCTION projects such an unfixed node from the initial $?Ty(t)$ node which bears a requirement for an expression of type *e* to be found at that node:¹³

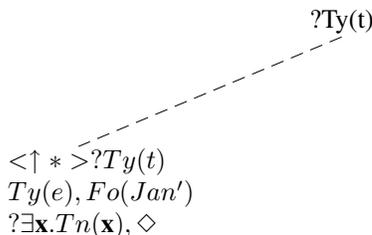
$$(26) \frac{\{...\{Tn(a), \dots ?Ty(t), \diamond\}\}}{\{...\{Tn(a), \dots ?Ty(t)\}, \{\langle \uparrow * \rangle Tn(a), ?\exists x.Tn(x), \dots, ?Ty(e), \diamond\}\}}$$

The effect of the rule is shown schematically below:



An NP can be parsed on that unfixed node satisfying the type *e* requirement:

- (28) Parsing *ton Jani* 'John' on an unfixed node¹⁴



¹²I will present two of the various variants of the ADJUNCTION rule here. Additional *ADJUNCTION rules will be introduced later on if needed.

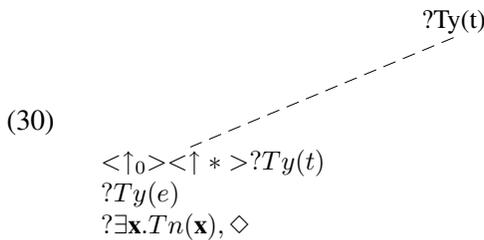
¹³The kleene star operator is a way to encode underspecification in the modal language. $\langle \uparrow * \rangle X$ reads as: X holds at a node above me or at the current node. Using the opposite modality, i.e the daughter modality, the kleene star denotes the notion of dominance plus reflexivity. The pure notion of dominance is encoded by means of the kleene plus operator (+). In that respect $\langle \uparrow + \rangle X$ reads as: There is a node above me where X holds. We will see later on how we will exploit both of the operators in our analysis.

¹⁴I ignore determiners for the moment.

The $\exists x.Tn(x)$ restriction will ensure that the node must be fixed at a later stage in the parse.¹⁵ The underspecified relation $\langle \uparrow * \rangle ?Ty(t)$ will enable the NP to be parsed in different structural positions. The *ADJUNCTION rule is a natural way to encode this intuition. The *ADJUNCTION rule will account for long scrambling cases as well, since the tree modality used does not restrict the full NP to apply in the local domain. A variant of *ADJUNCTION however, *LOCAL ADJUNCTION will do just that, i.e. it will restricts the potential fixing sites of the node to local nodes. The rule is shown below:

$$(29) \frac{\{...\{Tn(a), \dots, ?Ty(t), \diamond\}...\}}{\{...\{Tn(a), ?Ty(t), \dots\} \dots \{ \langle \uparrow_0 \rangle \langle \uparrow * \rangle Tn(a), ?\exists x.Tn(x), \dots, ?Ty(e), \diamond \} \dots \}}$$

The effect of the rule in tree notation is the following:



Notice that the modality has changed from $\langle \uparrow * \rangle$ to $\langle \uparrow_0 \rangle \langle \uparrow_1^* \rangle$. This will ensure that the NP in question is parsed in the local propositional domain.¹⁶ The two rules are used for long and short distance scrambling effects respectively. We will see later on the relevance of these rules with respect to clitics.

While the *ADJUNCTION rules involve the creation of an unfixed node that has a requirement for a specified treenode address in the tree under construction to be found, LINK structures involve the construction of a second tree structure independently of the initial one, which however posits a requirement for a shared term between the two trees. In order for LINK structures to be modelled, we need to introduce two new modal operators, $\langle L \rangle$ and $\langle L^{-1} \rangle$. The former refers to a tree structure which is linked, as it is shown schematically in (33), by means of an arrow, with the current node, while the latter refers to that node. LINK rules are also a family of rules, sharing the characteristics just mentioned. For demonstration purposes we will present one of them. The latter comes in the form of two rules, the rules of TOPIC STRUCTURE INTRODUCTION and TOPIC STRUCTURE REQUIREMENT¹⁷ respectively. These two rules are used by Cann et al. (2005) to account for Hanging Topic Left dislocation structures(HTLD). The first rule effectively creates a LINK transition between the initial node and a top node with a *type e* requirement. The rule is shown below:

$$(31) \frac{\{\{Tn(0), ?Ty(t), \diamond\}\}}{\{\{Tn(0), ?Ty(t)\}, \{ \langle L \rangle Tn(0), ?Ty(e), \diamond \}$$

Notice that the above rule does not mention anything about a shared term. That is because there is no shared term at the time the rule applies. The requirement for a shared term is introduced via the second rule the rule of TOPIC STRUCTURE REQUIREMENT shown below. This rule applies as soon as the dislocated argument is parsed (*as for Mary* in our example):¹⁸

¹⁵ $\exists x.Tn(x)$ reads as: There is a requirement for a proper treenode address (fixed) to be found. If the latter does not happen, then the parse cannot be completed as at least one outstanding requirement will exist in the tree.

¹⁶Assuming that all argument nodes are the 0 nodes and an additional propositional domain will involve a type t in one of the argument nodes, this rule will exclude cases where the NP is associated with an argument in the next propositional domain.

¹⁷The prototypical LINK rule is the rule of LINK *ADJUNCTION used by Cann et al. (2005) to account for relative clauses. We choose to present the TOPIC STRUCTURE INTRODUCTION and TOPIC STRUCTURE REQUIREMENT rules instead, which are basically a variant of the prototypical LINK rule. The interested reader is referred to Cann et al. (2005: 85-94) for the prototypical LINK rule.

¹⁸The D modality stands for the downward modality that encodes the kleene star operator and can furthermore extend over LINK structures In that respect D is defined as $D = \{\downarrow_0, \downarrow_1, \downarrow, \downarrow^*, L\}$.

$$(32) \frac{\{\{Tn(0), ?Ty(t)\}, \{< L > Tn(0), Fo(a), Ty(e), \diamond\}\}}{\{\{Tn(0), ?Ty(t), ? < D > Fo(a), \diamond\}, \{< L > Tn(0), Fo(a), Ty(e)\}\}}$$

After the introduction of these two rules, we get the following:

$$(33) \begin{array}{l} \langle L \rangle Tn(0) \\ Fo(Mary') \\ Ty(e) \end{array} \quad \langle L^{-1} \rangle Tn(n), ?Ty(t), ?\langle D \rangle Fo(Mary')$$

In an HTLD construction, the dislocated NP will be parsed in the node where the LINK begins. After TOPIC STRUCTURE REQUIREMENT has applied, a requirement for the same Formula value provided in the first tree will be put in the LINKed tree. This will ensure that a copy of the formula *Mary* will also be provided by the linguistic string, for example a resumptive pronoun in English. There are a number of important things with respect to the general LINK rule, but we won't discuss them here since these are not relevant to the scope of the paper. The interested reader is however referred to Kempson et al. (2001) and Cann et al. (2005) for more information on the various LINK rules.

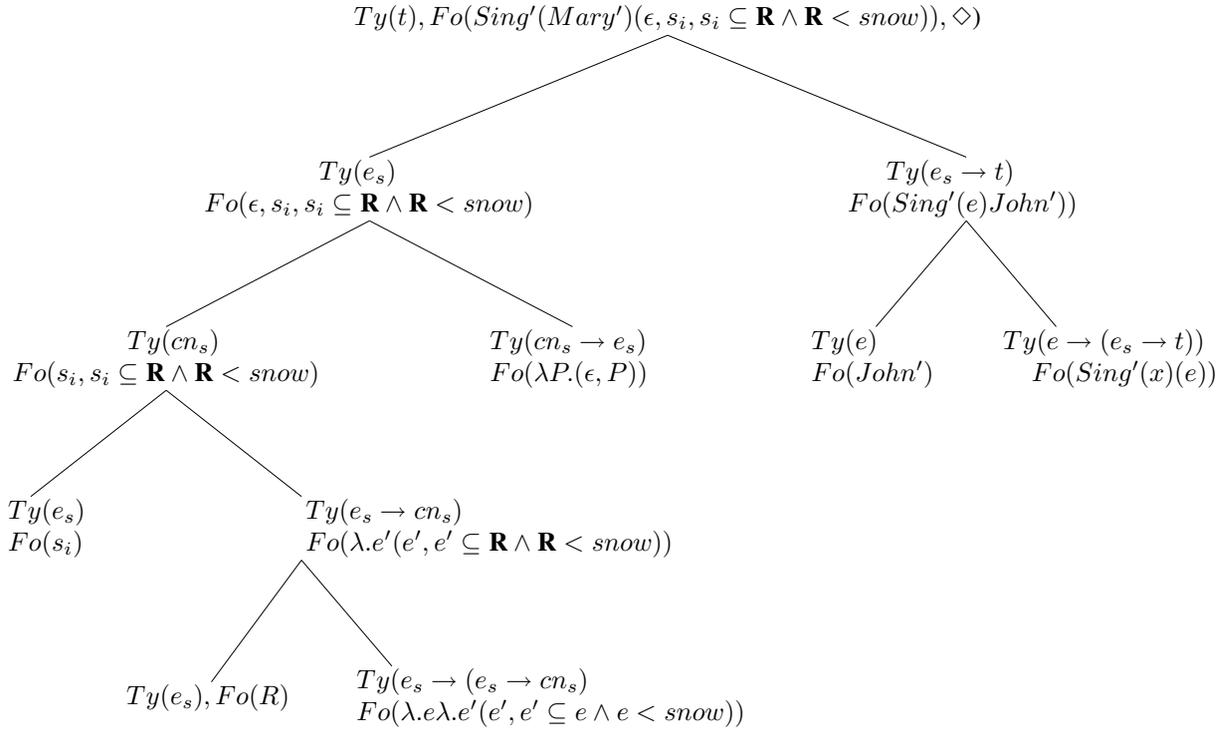
3.3 Newest developments in DS - Tense and Aspect

In both Kempson et al. (2001) and Cann et al. (2005), no attempt to address tense or aspect is made. Tense is encoded as a diacritic in Cann et al (2005) (e.g. Tns(Past)), noting that a proper analysis of tense is pending in the framework. Recent advances within DS however assume a treatment of tense based on the introduction of an explicit situation argument, introducing a situation in which the formula value the proposition expresses will be true.¹⁹ This situation argument node is then the locus of all tense and aspect information. In order to for this situation argument node to be represented two additional nodes are added to the standard DS propositional spine. A node standing for the situation argument, which is assumed to be of type e_s (with s standing for situation),²⁰ and a functor node of type $e_s \rightarrow t$. The situation argument node in line with quantified NPs (see Kempson et al., 2001; Cann et al., 2005 among others) is assumed to involve complex structure. The example below shows the structure assumed in Cann (forthcoming) after the complete parse of the sentence like 'Mary sang':

(34) Parsing *Mary sang*

¹⁹The situation argument node in DS was first introduced in Gregoromichelaki (2006).

²⁰In line with Gregoromichelaki (2006: 196) we assume that values in D_{Ty} involve $Ty(e)$ as a general type with subtypes e_i for individuals, e_s for situations etc. Cann (forthcoming) uses the notation e_w which is nothing more than a notational variant.



In the above structure, the intransitive verb *sing* is taken to be a two-place predicate, subcategorizing for both a subject and an event/situation argument. Let us see what the complex situation argument does. In the lowest e_s node, the reference time metavariable \mathbf{R} is introduced. This will combine with the semantic specifications given for the past tense in the lowest functor node ($Fo(\lambda.e\lambda.e'(e', e' \subseteq e \wedge e < snow))$), to return a formula value in which the first lambda bound variable (e) is substituted by \mathbf{R} ($Fo(\lambda.e'(e', e' \subseteq \mathbf{R} \wedge \mathbf{R} < snow))$). This new formula states that the remaining lambda bound variable e' is contained within or holds at \mathbf{R} ($e' \subseteq \mathbf{R}$) and that the reference time precedes the utterance time ($\mathbf{R} < snow$). In the intermediate e_s node, a situation s_i is introduced. This situation will substitute the remaining lambda bound variable (e') to return the formula value $Fo(s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < snow)$, in effect providing a situation that will bear the given tense/aspect specifications. The last step involves quantifying over the last formula we have obtained. In the example above, the situation with the given specifications is existentially quantified to return a formula value which roughly states that a past situation exists ($Fo(\epsilon, s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < snow)$). The last step involves substituting this last formula, for the lambda bound e in the formula value for *sing* to get the well-formed type t formula $Ty(t), Fo(Sing'(John'))(\epsilon, s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < snow)$.

Tense/aspect information are assumed to be projected mainly from verbs, both auxiliary and content verbs. However, a number of other elements can be taken to provide such information, like modality/tense particles/markers or even negation markers. With this said, I stop the discussion on the newest developments on tense/aspect in DS, noting that relevant additional details will be introduced if needed and directing the interested reader to Cann (forthcoming) for a detailed discussion of this approach.

4. A dynamic account of Clitic Climbing

4.1 Clitics in DS

A number of approaches have been proposed for clitics in DS (Bouzouita 2008a, 2008b; Chatzikiyriakidis 2009a, 2009b, forthcoming; Kempson & Chatzikiyriakidis 2009). In all these analyses, positioning restrictions are defined as restrictions on the current parse state, while the actions projected by the clitic vary depending on the level of underspecification involved in each case. For example, 1st/2nd person accusative clitics in Spanish have been proposed (Kempson & Cann, 2007; Kempson & Chatzikiyriakidis, 2009) to project locally unfixed nodes, a proposal largely motivated by the syncretism found in these clitic forms. On

the other hand, 3rd person accusative clitics in the same language are treated as projecting fixed structure, in effect building and decorating the direct object node. There are a number of interesting predictions that such a proposal makes, especially with respect to the PCC but this is something that will not bother us in this paper. The interested reader is however directed to Kempson and Cann (2007) and Kempson and Chatzikyriakidis (2009) for an analysis of the PCC in DS. In this paper I will use the Italian 3rd person accusative neuter clitic *to* as a role model and I will not deal with clitic clusters or person case restrictions. The role language used will be Italian but the same account will be easily extendable to other CC languages as well (see Chatzikyriakidis 2010 for a DS analysis of optional climbing.). As already said, positioning restrictions are defined as restrictions on the current parse state. For example the trigger shown below, effectively captures proclisis associated with clitics in Italian:²¹

(35) Trigger ensuring proclisis

```

IF      ?Ty(t)
THEN   IF      [↓1+?Ty(x)
           THEN ...
ELSE   abort

```

The above reads as follows: If you are in a node with a type *t* requirement, then if all functor nodes (nodes with a *T_n* address) below that node bear a type requirement, then the clitic can be parsed. This means that no verbal type has been parsed yet, since in case it did we would have at least one fixed node.²² The next step is to define the actual actions projected by the clitic. Given a fixed node analysis for 3rd person accusative clitics, *to* will basically build the direct object node, project a type value and a formula metavariable on that node and return the pointer to the most local type *t* node above it (*gofirst*(?Ty(*t*))). Furthermore, an additional trigger is added, specified as a disjunction in the embedded IF part of the algorithm (OR), positing that the clitic can also be parsed in case an imperatival feature is present in the type *t* requiring node:²³

(36) Lexical entry for the third person accusative clitic *to*²⁴

```

IF      ?Ty(t)
THEN   IF      [↓1+?Ty(x)
           OR
           IF      +IMP
           THEN   make(⟨↓1⟩); go(⟨↓1⟩);
                 make(⟨↓0⟩); go(⟨↓0⟩)
                 put(Ty(e), Fo(Ux), ?∃x.Fo(x), gofirst(?Ty(t)))
ELSE   Abort

```

Having sketched the way clitics are treated in DS, it is time to move to the actual analysis of CC in DS.

4.2 The analysis

The problem posed by CC in GSG, given our account of clitics and assuming a biclausal structure is involved in CC constructions, is that the clitic ends up projecting a type *e* value in the direct object node

²¹The same trigger is a general proclitic trigger and as such will work for languages with similar clitic positioning restrictions (Spanish, Italian, French to name a few.)

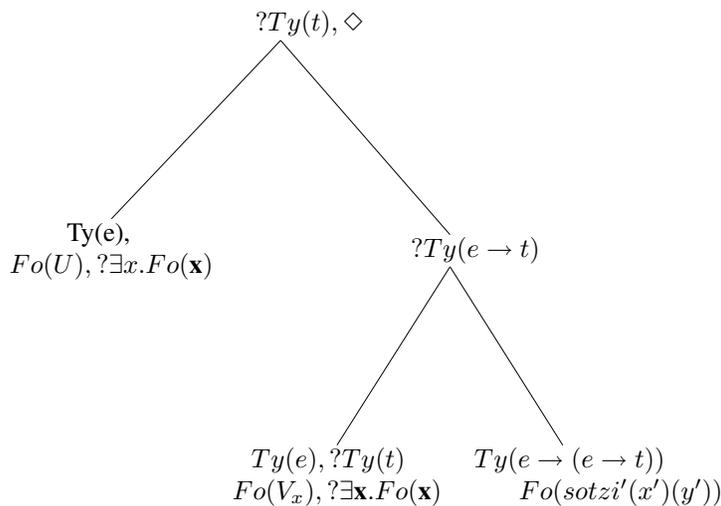
²²Note that such an entry will also correctly predict *to* to be possible after dative clitics in clitic clusters like *mu to*. Assuming that *mu* will project an unfixed node (following Kempson and Cann, 2007; Kempson and Chatzikyriakidis, 2009), the clitic *to* can be parsed.

²³Note that the feature +IMP are used as diacritics and do not by any means constitute any serious attempt to give an analysis of imperatives and infinitives in DS.

²⁴The subscript *x* in the formula metavariable *U* stands for the restrictions on metavariable update that the clitic will bear. These will not be specified in this paper but will have to be assumed in a more complete analysis to prevent overgeneration, e.g. avoiding a situation where *it* is updated by a formula value specified as female.

projected by the modal that however bears a type t requirement, standing for the type the infinitival clause will receive. The partial tree after the modal *sotzo*, ‘can’, in *to sotzo vorasi*, ‘I can buy it’, is parsed is shown below:

(37) Parsing *sotzi* in *to sotzi vorasi*



Notice that both a type e value and type t requirement exist in the direct object node. Such a partial tree cannot lead to a well-formed parse. The reason is simple. Satisfying the type t requirement will lead to a situation where the node carries two distinct, incompatible type values, something obviously not allowed by the system. Leaving the requirement unsatisfied will not do any better, since an outstanding requirement will always exist in the tree. Since a complete tree representing a successful parse, as already mentioned in the introduction to the framework, must not have any outstanding requirements, the parse will never be completed in case the type t requirement does not get satisfied. This is the situation we get assuming climbing inducing verbs are parsed like regular verbal complement verbs, i.e. when a biclausal structure is assumed to be involved in CC constructions. However, CC has been argued convincingly to involve a monoclausal rather than a biclausal structure (see Cinque 2006 for an extensive discussion). The next question that comes to mind is how DS can capture this fact, or more practically how such a thing is formalizable in DS. The answer that I will propose, basically assumes that climbing inducing verbs behave like auxiliary verbs in that they do not project a verbal functor type, but rather project their semantics inside a complex situation argument node. According to Ronnie Cann (forthcoming) English auxiliaries are taken to project tense and aspect in the situation argument node complex, while they further project a type value and a formula metavariable in the predicate node (011 node). The same assumptions can be used to analyze auxiliaries *eço/ixa* ‘have/had’ in SMG. In SMG these two auxiliaries are used to form the present and past perfect respectively. In the presence of a clitic in perfect constructions, climbing of the clitic before the auxiliary is obligatory:

(38)

- a. *To* *exo* *δesi*
it.CL-ACC have.1SG tied.PST-PRTCPL
- b. **Exo* *to* *δesi*
have.1SG it.CL-ACC tied.PST-PRTCPL
‘I have it tied.’ [SMG]

In line with Cann’s (forthcoming) analysis for auxiliaries in English, I assume that in SMG auxiliary *eço*, projects all the relevant semantic information in the complex situation argument node and thus no verbal type value needs to be projected from the auxiliary. However, unlike English which projects a formula

metavariable ($Fo(U)$) and a type value containing a metavariable ($Ty(U \rightarrow (e_s \rightarrow t))$) in the 011 node, SMG auxiliaries do not project the 011 at all. They do however project the subject node along with a formula metavariable and a type value. The reason for the latter move is that subject agreement in SMG is encoded in the auxiliary. Thus, all the relevant information that make pro-drop possible in SMG should be encoded in there. They eventually as in the English case leave the pointer at the situation functor node. The lexical entry plus the result of parsing *exo*, ‘have’, is shown below:

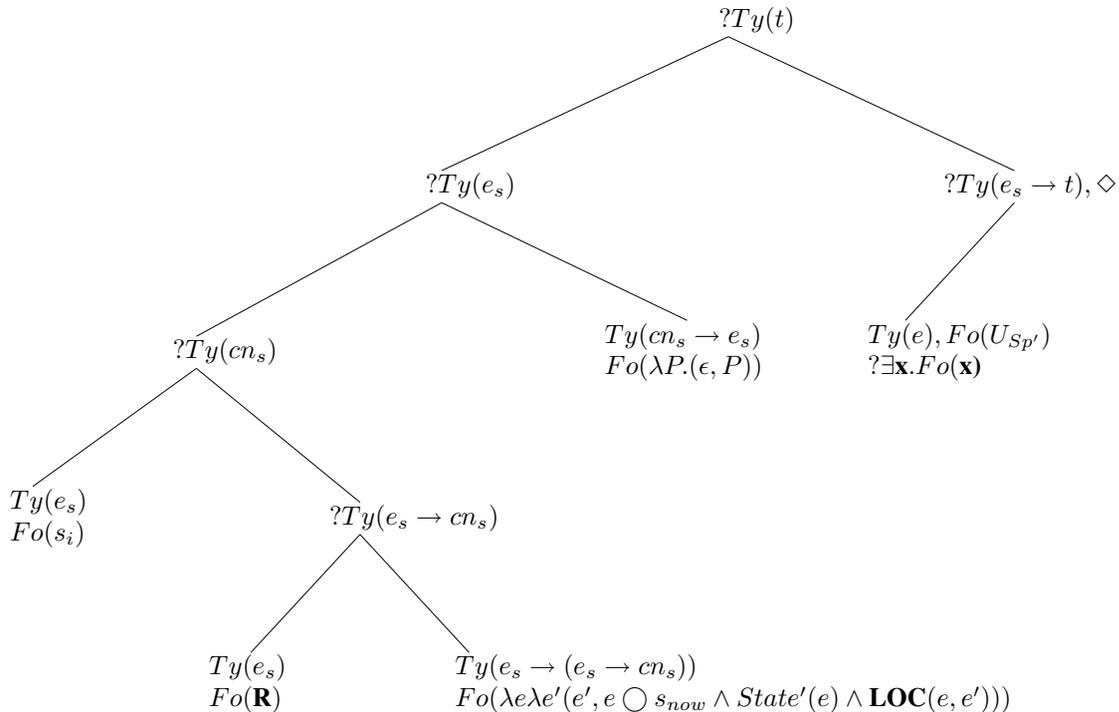
(39) Lexical entry for *exo* ‘have.1SG’ in SMG

```

IF      ?Ty(t)
THEN   make(⟨↓0⟩); go(⟨↓0⟩); put(?Ty(es))
       make(⟨↓1⟩); go(⟨↓1⟩); put(Ty(cn → es), Fo(λP.(ε, P)))
       go(⟨↑1⟩); make(⟨↓0⟩); go(⟨↓0⟩); put(?Ty(cns))
       make(⟨↓0⟩); go(⟨↓0⟩); put(Ty(es), freshput(s))
       make(⟨↓0⟩); go(⟨↓1⟩); put(?Ty(es → cn))
       make(⟨↓0⟩); go(⟨↓0⟩); put(Ty(es, Fo(R)))
       go(⟨↑0⟩); make(⟨↓1⟩); go(⟨↓1⟩); put(Ty(es → (es → cn)))
       put(Fo(λeλe'(e', e ○ snow ∧ State'(e) ∧ LOC(e, e'))))
       go(⟨↑1⟩⟨↑1⟩⟨↑0⟩⟨↑0⟩); make(⟨↓0⟩); go(⟨↓0⟩); put(Ty(e), Fo(USpeaker'), ?∃x.Fo(x))
       go(⟨↑1⟩)make(⟨↓1⟩); go(⟨↓1⟩); put(?Ty(es → t))
ELSE   abort

```

(40) The effect of parsing *exo*



The auxiliary introduces both the reference time R and a fresh situation s_i . Furthermore, it introduces the tense/aspect specifications for the present perfect ($Fo(\lambda e\lambda e'(e', e \circ s_{now} \wedge State'(e) \wedge \mathbf{LOC}(e, e'))$), where \circ stands for the overlap relation while \mathbf{LOC} expresses an underspecified relation between the event and the reference points that enables the various perfect readings to be generated.²⁵ Additionally, the auxiliary projects the formula $Fo(\lambda P.(\epsilon, P))$ which will basically existentially quantify over the formula value that will emerge after combining the formula value of the intermediate e_s node with the formula value of the

²⁵See Cann (2009) for details and motivation as regards the \mathbf{LOC} feature.

?Ty($e_s \rightarrow cn$) requiring node. The rest of the structure projected is rather straightforward. Assuming this entry for auxiliaries and furthermore the entries we have given for clitics in SMG, climbing is predicted to be the only option with auxiliaries. Let us see why. Assuming a clitic has been parsed first, the pointer is left at the type t requiring node. This node is then the trigger for parsing the auxiliary (see lexical entry 39) and the rest follows naturally. On the other hand, assuming that the auxiliary has been parsed first the pointer is left at the situation functor node. In case a clitic comes into parse, the parse will abort, since the initial triggering point I have given for clitics, i.e. a type t requiring node, will not be satisfied given that the pointer will be at the situation functor node (?Ty($e_s \rightarrow t$)). Notice that the pointer cannot move up via COMPLETION, since no type or formula will be satisfied in the situation functor node in parsing an auxiliary. In that respect, the only option is for the clitic to appear proclitic to the auxiliary according to fact. Notice furthermore, that the auxiliary *exo*, ‘have’, in SMG unlike its English counterpart (Cann, forthcoming) does not project a type value and a formula metavariable in the 011 node. This is because assuming a type value and a formula metavariable are projected by the auxiliary in the 011 node, VP ellipsis will be predicted to be possible with auxiliaries in SMG contrary to fact. Given the projection of a formula metavariable in the 011 node by the auxiliary, substituting this formula metavariable with a value provided by the context (which is always a possibility for metavariables), can give us a well formed sentence. This works nicely for English but will fallaciously predict VP ellipsis with auxiliaries to be possible in SMG. The examples below show the relevant facts for English and SMG:

(41) *Have you hit John? Yes, I have*

(42) *Exis xtipisi ton Jani? *Ne, eço (SMG)*
 have hit the_{acc} John_{acc} yes have
 ‘Have you hit John? Yes, I have.’

So far, so good. The question how is this account relevant to restructuring verb climbing? The answer is that an analogous account can straightforwardly be put forth for restructuring verbs by simply making the following two assumptions: a) Climbing inducing verbs do not project any verbal functor type and b) The semantics of restructuring verbs are captured in the complex situation argument node similarly to auxiliaries. Let us illustrate this claim using the GSG verb *sotzo* ‘can’. As already said, following Cann (forthcoming) I take modals to behave like auxiliaries in that they project their semantics in the situation argument node rather than projecting a verbal type. However, remember that modals are content verbs and more than tense and aspect information will be needed to capture their semantics. Fortunately, there is a way in which this can be done. Remember that aspect and tense information were introduced in the Ty($e_s \rightarrow (e_s \rightarrow cn)$) node of the complex situation argument node and percolate up to the Ty($e_s \rightarrow cn$), where they combine with the situation Fo(s_i) provided by the intermediate e_s node. Now, assuming that this situation (Fo(s_i)) can also be evaluated with respect to possible worlds we immediately get a solution to our problem. The only thing that we will have to further assume is that a ‘world’ parameter is also projected as part of a complex argument involving a situation and a world parameter, both independently quantified by the right quantifier in each case. The assumption I am going to make is that modal verbs project such a complex situation argument involving a situation and a world parameter. Then, the next step is the use of possible world semantics. For example, the lexical entry for *must* can be seen as specifying that the proposition expressed by the infinitive plus its arguments is true in all contextually given possible worlds accessible from the default world.²⁶ The same can be argued to be the case for the ability modal *sotzo* ‘can’, the difference being that the domain of quantification in this case is ability worlds, a subset of the set of possible worlds rather than just possible worlds. *Sotzo* under such an approach, will project a complex Fo value in the internal e_s node, encoding both a situation and a world parameter (Fo(s_i, w_i)). This world parameter must be a member of the set of ability

²⁶A fact already suggested to me by Ronnie Cann (p.c).

contexts which in turn are a subset of the set of contextually accessible worlds. The next thing we need to take care of is the form of quantification quantifying over these possible worlds. Since we are dealing with the set of all ability contexts, what we need is universal quantification. In that respect, we posit a tau term, instead of an epsilon term to capture the universal quantification properties of the world parameter projected by *sotzo*. Furthermore, tense/aspect specifications are also going to be included. Under the account just sketched, the only difference between modals and auxiliaries is that the former introduce a complex situation argument including both a situation and a world parameter in the intermediate e_s node, whereas the latter projects only a situation parameter. One further difference between the two is the node where the pointer is assumed to be left. We have seen that the pointer is left at the $e_s \rightarrow t$ node after an auxiliary is parsed. However, leaving the pointer at the same node in the case of restructuring verbs will basically predict that infinitives have two distinct parsing triggers, a type t and a type $e_s \rightarrow t$ requiring trigger. The type t requiring trigger is independently needed for constructions where the infinitive functions as the complement of a regular complement taking verb. In that case, and assuming that the complement taking verb will decorate the direct object node with a type t requirement and will leave the pointer there with no other nodes existing below that node, the trigger for the infinitive must be the type t node. In order to avoid redundancy, I posit that the trigger for infinitives is a type t requiring node in all cases²⁷. Lastly, I further assume that restructuring verbs in GSG further project the predicate node, in contrast to SMG auxiliaries and similarly to English auxiliaries and modals. The reason for this is that VP ellipsis is possible with restructuring verbs in GSG:

- (43) *To sotzi vorasi? Ne, sotzi*
 it.CL-ACC can.3SG buy.INF yes can.3SG
 ‘Can he buy it? Yes, he can.’

Assuming that *sotzi* projects a formula metavariable in the predicate node along with a type value, then given standard DS assumptions (Kempson et al., 2001; Cann et al., 2005) on metavariable update, this update can be done via the context in which case no overt input is needed.

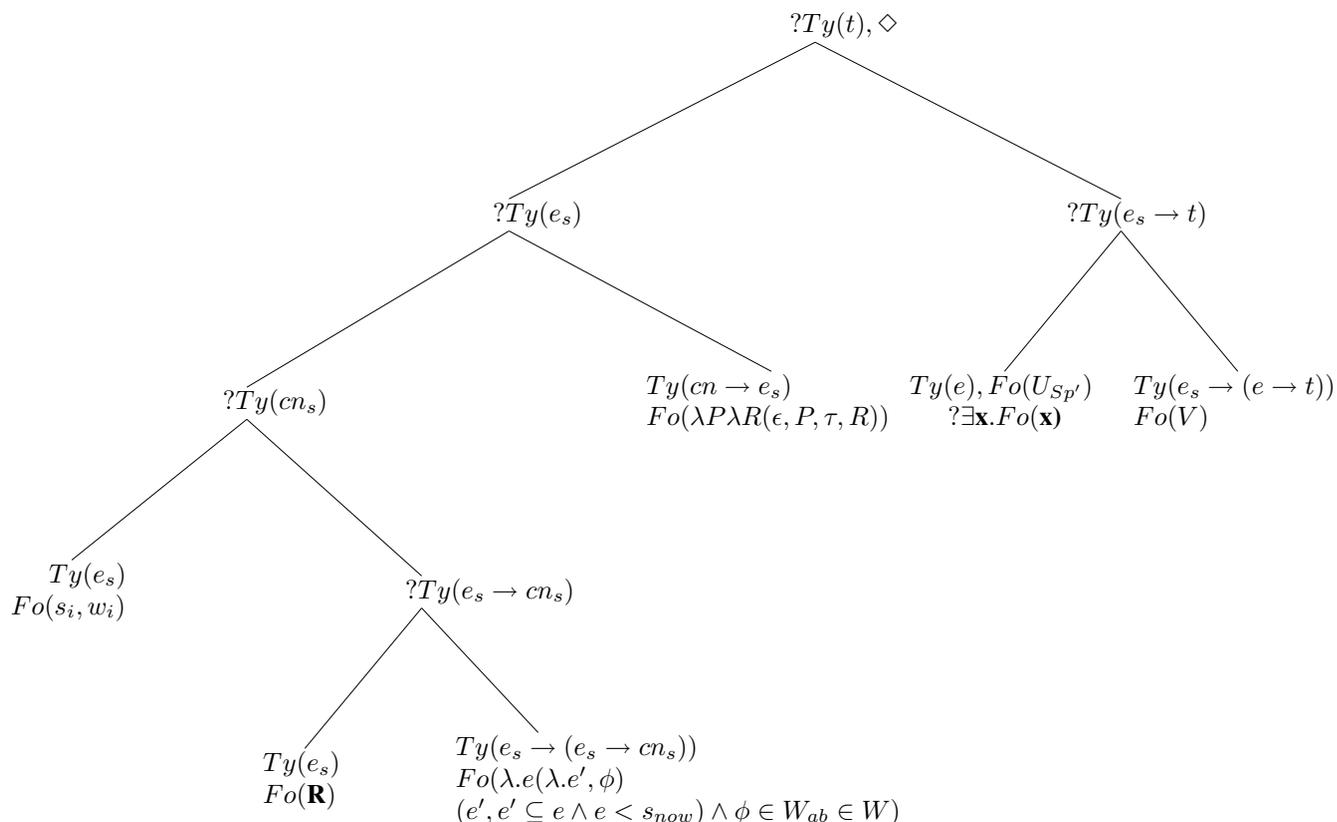
Putting all these assumptions together, one gets the lexical entry for ‘*sotzo*’ shown below:

- (44) Lexical entry for *sotzo*.1SG, ‘can’ in GSG
 IF ?Ty(t)
 THEN $\text{make}(\langle \downarrow_0 \rangle); \text{go}(\langle \downarrow_0 \rangle); \text{put}(?Ty(e_s))$
 $\text{make}(\langle \downarrow_1 \rangle); \text{go}(\langle \downarrow_1 \rangle); \text{put}(Ty(cn \rightarrow e_s), Fo(\lambda P \lambda R(\epsilon, P, \tau, R)))$
 $\text{go}(\langle \uparrow_1 \rangle); \text{make}(\langle \downarrow_0 \rangle); \text{go}(\langle \downarrow_0 \rangle); \text{put}(?Ty(cn))$
 $\text{make}(\langle \downarrow_0 \rangle); \text{go}(\langle \downarrow_0 \rangle); \text{put}(Ty(e_s), \text{freshput}(w_i, s_i))$
 $\text{make}(\langle \downarrow_1 \rangle); \text{go}(\langle \downarrow_1 \rangle); \text{put}(?Ty(e_s \rightarrow cn))$
 $\text{make}(\langle \downarrow_0 \rangle); \text{go}(\langle \downarrow_0 \rangle); \text{put}(Ty(e_s), Fo(R))$
 $\text{go}(\langle \uparrow_0 \rangle); \text{make}(\langle \downarrow_1 \rangle); \text{go}(\langle \downarrow_1 \rangle); \text{put}(Fo(\lambda.e(\lambda.e', \phi)$
 $(e', e' \subseteq e \wedge e < snow) \wedge \phi \in W_{ab} \in W))$
 $\text{go}(\langle \uparrow_1 \rangle \langle \uparrow_0 \rangle \langle \uparrow_0 \rangle \langle \uparrow_0 \rangle); \text{make}(\langle \downarrow_1 \rangle); \text{go}(\langle \downarrow_1 \rangle); \text{put}(?Ty(e_s \rightarrow t)$
 $\text{make}(\langle \downarrow_0 \rangle); \text{go}(\langle \downarrow_0 \rangle); \text{put}(Ty(e), Fo(U_{Speaker'}), ?\exists \mathbf{x}. Fo(\mathbf{x}))$
 $\text{go}(\langle \uparrow_1 \rangle \langle \uparrow_0 \rangle \langle \uparrow_0 \rangle) \text{put}(Ty(e_s \rightarrow (e_s \rightarrow e_s)), Fo(V), \text{go} \text{first}(?Ty(t))$
 ELSE abort

The result of parsing *sotzo* is shown below:

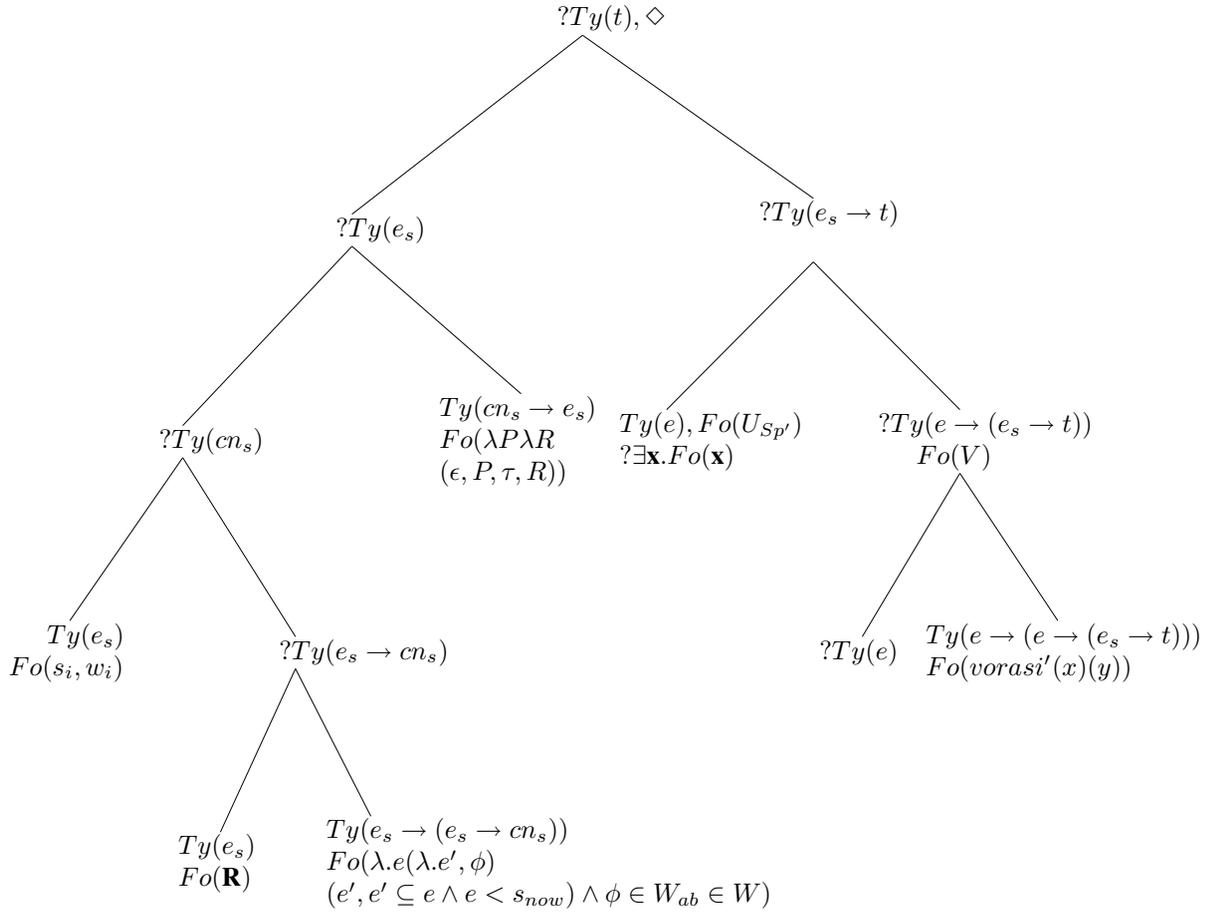
- (45) Parsing *sotzo*

²⁷A further welcomed result is that such treatment can further enable us to distinguish between infinitives and past participles without actually referring to any other of their properties. For example, under such an analysis, an infinitive will always be impossible after an auxiliary has been parsed first, since the pointer in that case will be at the situation functor node.



The intermediate $Ty(e_s)$ node has a complex formula value introducing both a situation and a world parameter ($Fo(s_i, w_i)$). The lowest functor node on the other hand ($Ty(e_s \rightarrow (e_s \rightarrow cn_s))$) contains three lambda bound variables ($Fo(\lambda.e(\lambda.e', \phi)(e', e' \subseteq e \wedge e = s_{now}) \wedge \phi \in W_{ab} \in W)$). The first variable (e) is to be substituted by the reference time metavariable \mathbf{R} . Then, the other two (e' and ϕ) stand for the two variables that will be substituted by the two parameters of the complex Fo value of the internal e_s node, s_i and w_i respectively. In that sense, e (\mathbf{R} after substitution) is taken to hold at a time e' , where time e' is the same as the utterance time s_{now} , while ϕ is taken to belong to the set of ability contexts, which in turn are a subset of the contextually accessible worlds W ($\phi \in W_{ab} \in W$). Then, at the 001 node, the form of quantification for each of the arguments is introduced. The situation parameter is associated with existential while the world parameter with universal quantification ($Fo(\lambda P \lambda R.(\epsilon, P, \tau, R))$). The subject node is further projected (010 node) and a type value and a formula metavariable are posited in the same node. Lastly, the predicate node is also built and decorated with a type value and a formula metavariable. The pointer is left at the type t requiring node. Given the structure projected by *sotzo* in (45), the clitic cannot be parsed after *sotzo* has already done so. The trigger of the accusative clitic will not be satisfied because a functor node will bear a type value (the 011 node), while the trigger for genitive clitics will not be satisfied given that a number of fixed nodes will exist after parsing *sotzo*. Parsing of the clitic after the infinitive is not possible for the same reasons. The infinitive builds the 0111 node and projects a verbal type and a formula value on that same node. Furthermore, it also builds the direct object node and decorates it with a type e requirement. Lastly, it returns the pointer to the type t requiring node:

(46) Parsing *vorasi* ‘to buy’ in **sotzo vorasi to* ‘I can buy it’

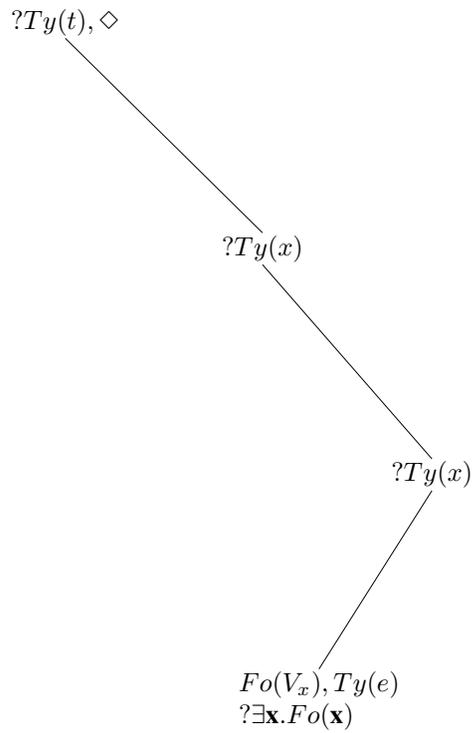


It is at that point that the clitic comes into parse. The lexical entry for the clitic *to* ‘it’ specifies that no functor node with a type value must exist in order for the parsing process to proceed. However, this is not true in the tree above, since two functor nodes with type values exist (the 011 and 0111 nodes). Thus, CC is the only option in the presence of *sotzo* in GSG. In a CC case like the one shown below, the clitic is parsed first, building and decorating the direct object node with a type value and a formula metavariable:

(47) *To sotzo vorasi*
 it.CL-ACC can.1SG buy.INF
 ‘I can buy it.’

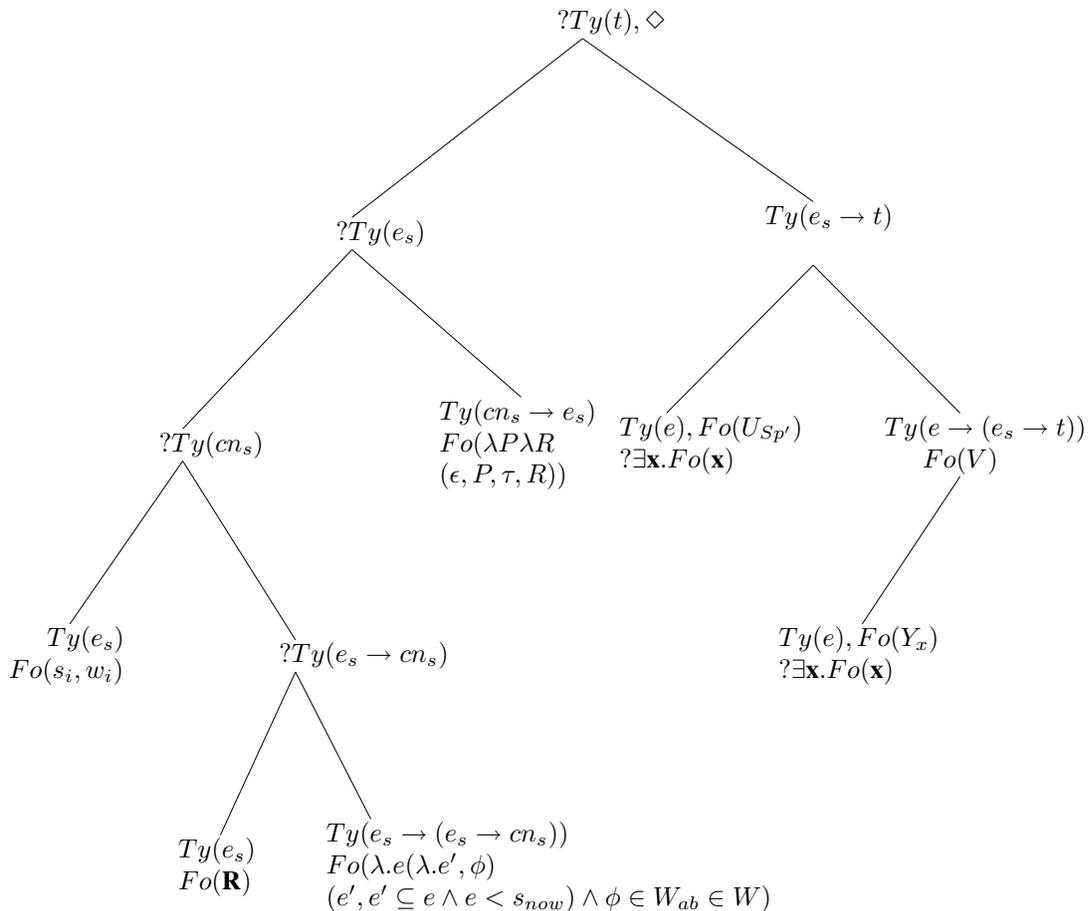
(48) Parsing *to* ‘it’ in *to sotzo vorasi* ‘I can buy it.’

Clitic Climbing in GSG: a dynamic account



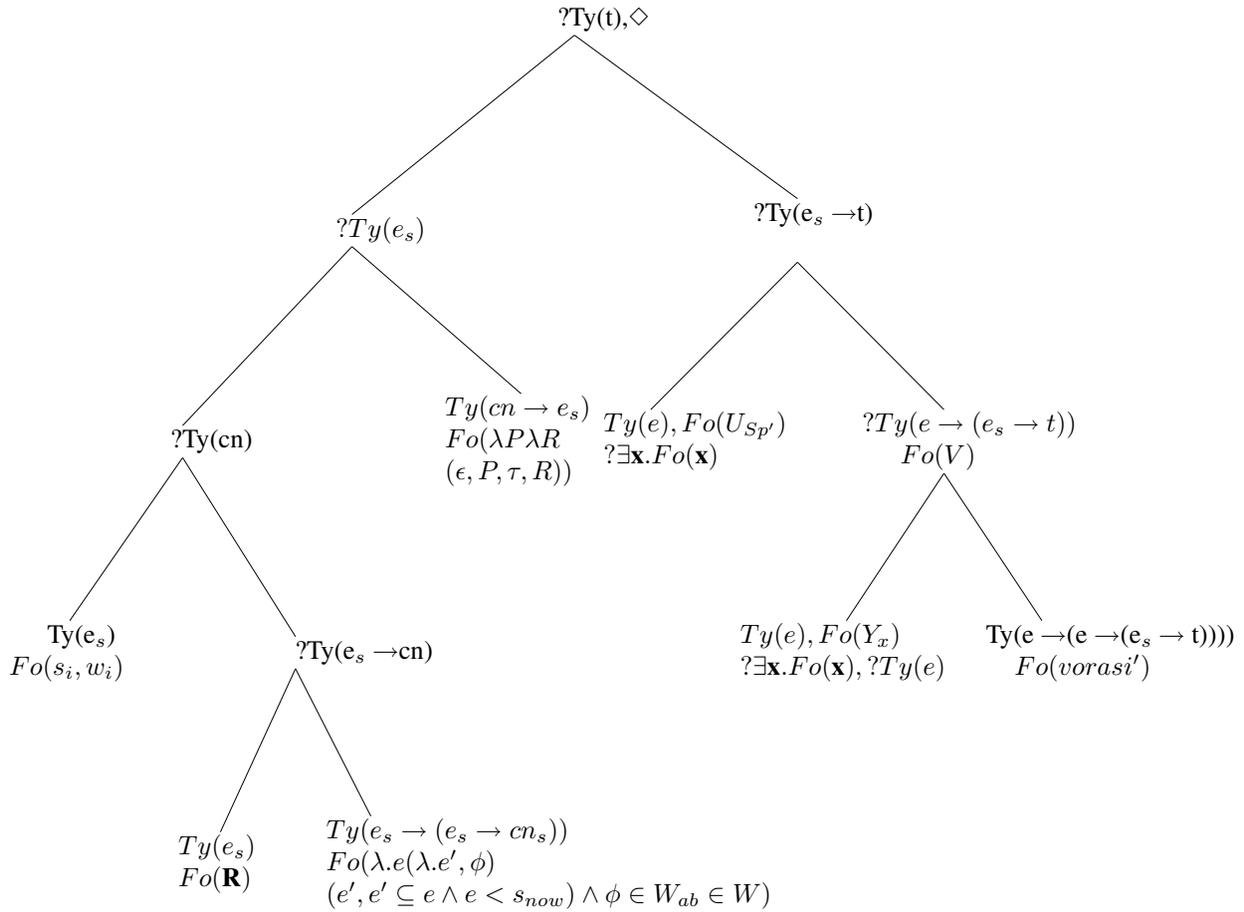
With the pointer being at the type t requiring node, *sotzi* comes into parse:

(49) Parsing *sotzo* ‘can’ in *to sotzo vorasi* ‘I can buy it’



The pointer is again at the type t requiring node. The infinitive comes into parse, projecting a type plus a formula value in the 0111 node. It further builds the direct object node and decorates it with type e requirement:

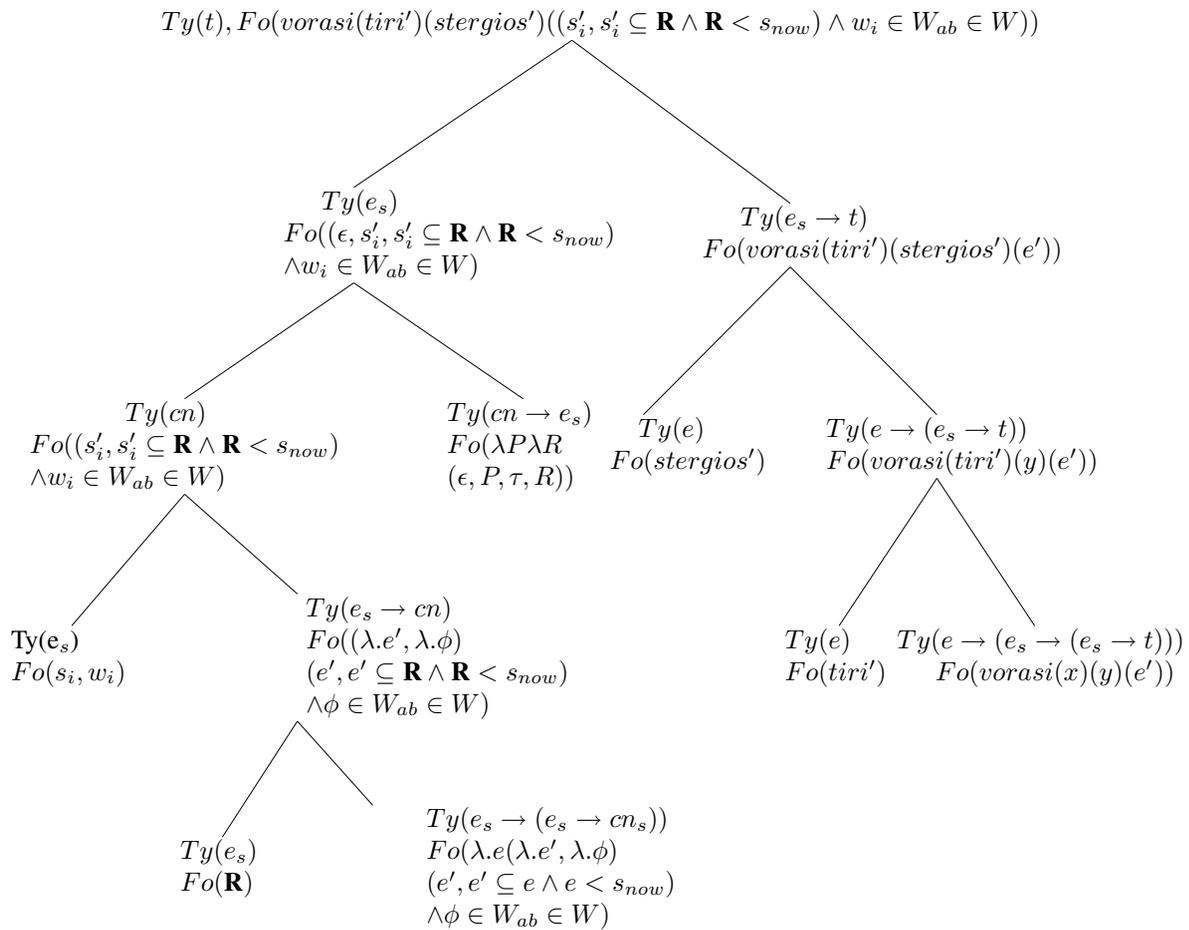
(50) Parsing *vorasi* ‘can’ in *to sotzo vorasi* ‘I can buy it’



At that point the rules of ELIMINATION, THINNING, COMPLETION and metavariable SUBSTITUTION will apply getting rid of any requirements that have been satisfied, combining formula and type values via functional application and modus ponens respectively and providing proper Fo values for the metavariables from the context. The result of all these processes is the well-formed parse shown below:

(51) Completing the parse

Clitic Climbing in GSG: a dynamic account



The difference between optional climbing languages on the one hand and GSG on the other is that in the latter, infinitives are incompatible with clitics whereas in optional climbing languages infinitives can host clitics. In that sense, the lexical entries in these languages will have to involve a trigger that will capture enclitic positioning with infinitives as well (see Chatzikyriakidis 2009, forthcoming, in preparation for a DS analysis of optional CC).

The two verbs climbing inducing verbs in GSG can form multiple climbing constructions, in which the clitic climbs across two verbs:

- (52) (T)o sotzo spiccetsi tse di avri
it.CL-ACC can finish.INF COMP see.INF tomorrow
'I can finish seeing it tomorrow.'
- (53) *Sotzo spiccetsi tse to di avri
can spiccetsi COMP it.CL-ACC see.INF tomorrow
'I can finish seeing it tomorrow.'
- (54) *Sotzo spiccetsi tse di to avri
can spiccetsi COMP see.INF it.CL-ACC tomorrow
'I can finish seeing it tomorrow.'

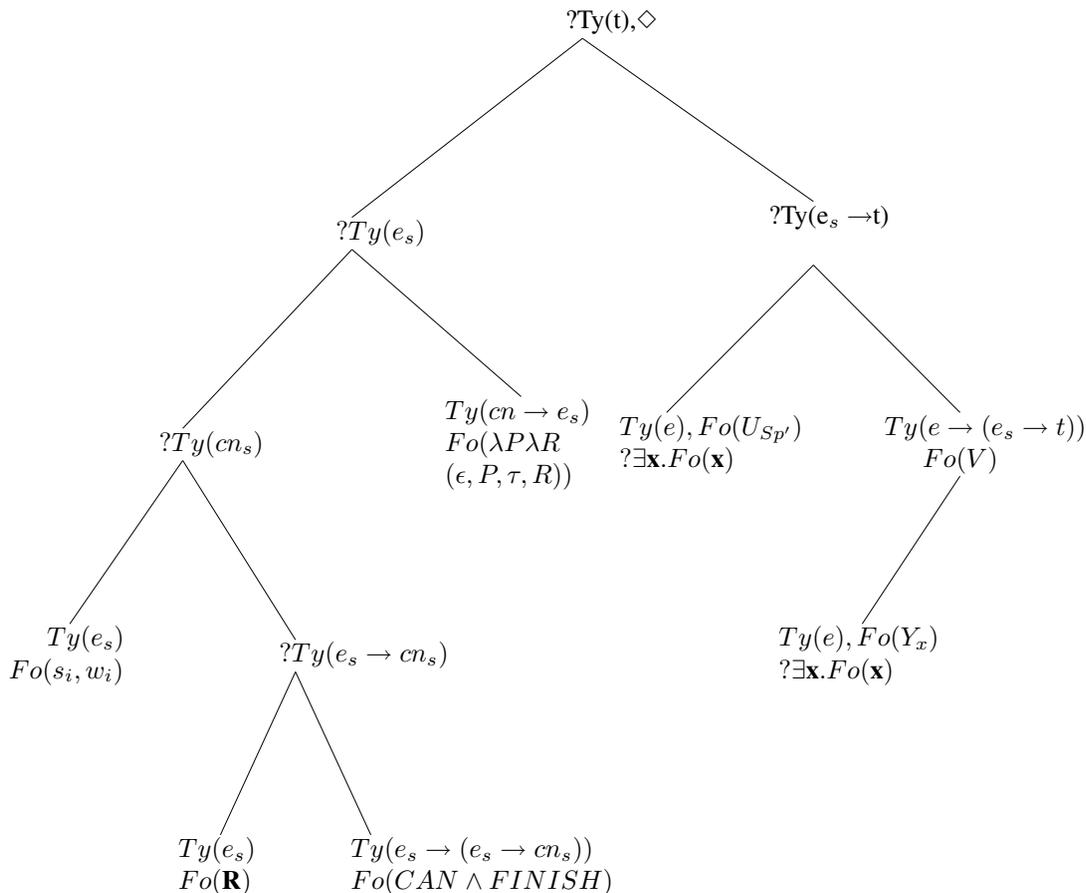
Climbing in the intermediate position is also banned in GSG, in contrast to languages like Italian where intermediate climbing is permitted:²⁸

²⁸A number of restrictions apply in case of multiple climbing in Italian. See Cardinaletti & Shlonsky (2004) for an extensive discussion on the issue.

- (55) **Sotzo to spiccetsi tse di avri*
 can it.CL-ACC finish.INF COMP see.INF tomorrow
 'I can finish seeing it tomorrow.'
- (56) **Sotzo spiccetsi to tse di avri*
 can finish.INF it.CL-ACC COMP see.INF tomorrow
 'I can finish seeing it tomorrow.'

The account sketched so far correctly predicts the above facts. Let us see why. The restructuring infinitive is assumed to be parsed in the same sense as *sotzo*, i.e. as encoding its semantics in the situation argument node. The difference between finite and infinitive restructuring verbs lies in the fact that infinitives in contrast to finite verbal forms will not introduce a freshput situation or world parameter in the internal e_s node but will rather depend on the situation or world already projected by the finite verb. Tense, aspect or world specifications will be projected on the relevant nodes. If these nodes already contain such specifications, a combination of the two specifications is done using a form of generalized conjunction in the sense of Cann (forthcoming), effectively combining two formulae of the same type. The example below illustrates the result of parsing *sotzo spiccetsi* 'can finish':²⁹

- (57) Parsing *sotzo spitsetsi*

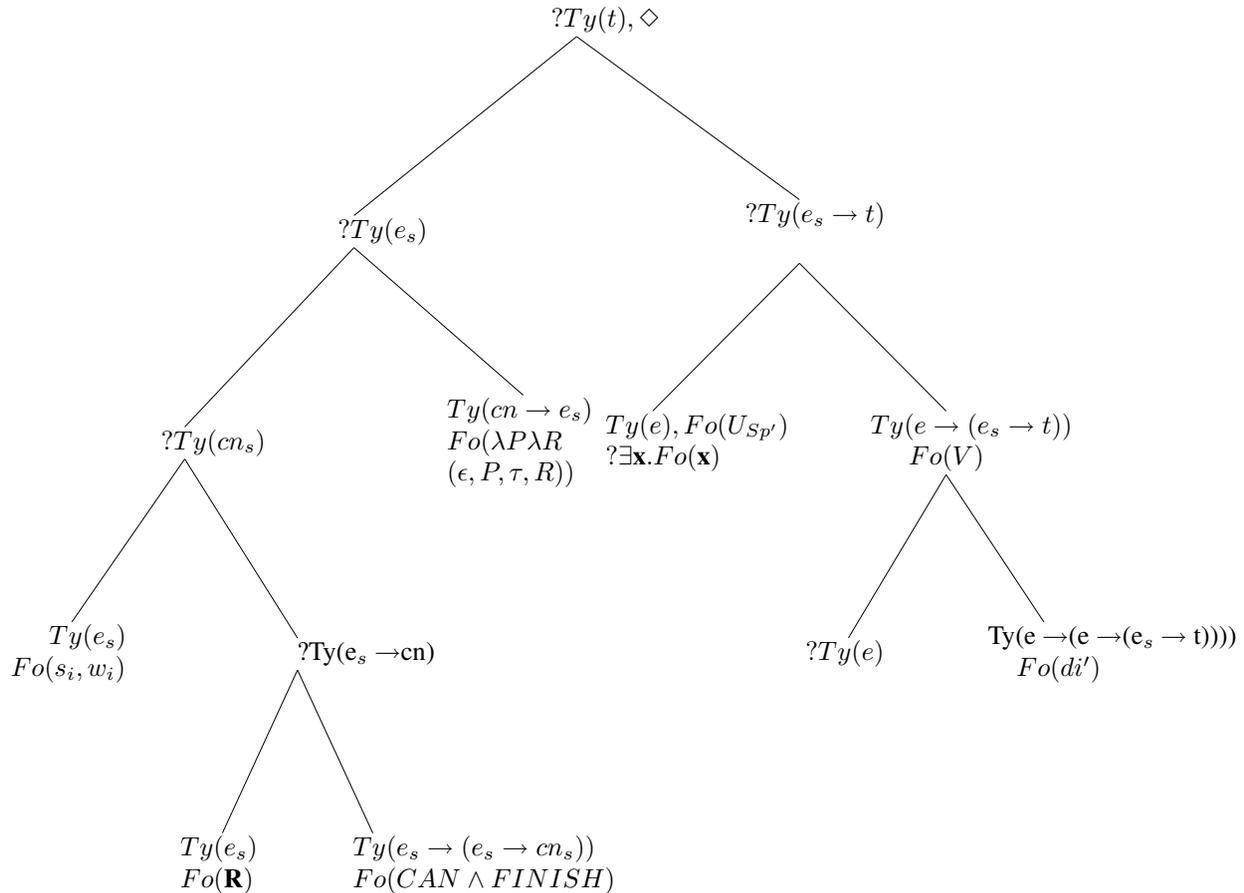


Example (53) is predicted to be ungrammatical, since the clitic trigger $[\downarrow_1^+]$?Ty(x) is again not satisfied (a verbal type and a number of fixed nodes exist). Thus the parse will abort. In (54) the clitic will come into

²⁹Note that the exact formal specifications of what both *sotzo* and *spitsetsi* contribute to the 0001 node are not shown. This is because the exact semantics have not been fully worked out yet in this case. The statement *CAN ∧ FINISH* is only a diacritic for what the exact semantics projected by *sotzo* and *spitsetsi* will be.

parse after the lexical infinitive will project its structure, i.e. the rest of the propositional template plus a verbal type in the 0111 node:

(58) After parsing *di* 'see.INF' in **sotzo spiccetsi tse di to avri*



The clitic cannot be parsed since in the above partial tree a functor type is present ($Ty(e \rightarrow (e \rightarrow (e_s \rightarrow t)))$), thus the triggering restrictions we have given for clitics do not get satisfied. The trigger for clitics ($[\downarrow_1^+]?Ty(x)$) will fail due to the presence of two functor types (in the 011 and 0111 nodes). Similarly in (55) and (56), parsing of a clitic is not possible, since again one functor type will be present (in the 011 node). The only option in that respect under our account is multiple CC in accordance with the facts. The account I have proposed correctly predicts the climbing facts in GSG. Furthermore, CC with auxiliary verbs in SMG is also accounted correctly within the same reasoning. There are a number of other welcoming results this account has to offer for a number of other phenomena found in CC languages, notably unavailability of negating infinitives in CC contexts³⁰ and auxiliary switch in Italian but these will not be discussed here for reasons of relevance. The reader interested in the way DS can account for these phenomena is directed to Chatzikyriakidis (2009, forthcoming).

5. Conclusion

In this paper, I have presented a first sketch of a DS account of CC in GSG. I have argued that the phenomenon of CC can receive a straightforward explanation and formalisation once one shifts into a dynamic perspective. In particular, I provided an analysis of restructuring verbs as auxiliary-like verbs based on the auxiliary analysis given by Cann (forthcoming). Under this analysis, restructuring verbs do not project a verbal type but rather project their semantics inside the complex situation argument node. This assumption

³⁰Such a fact is precluded independently for GSG, since infinitives cannot be negated in general in GSG or Calabrian Greek (Katsoyannou, 1995).

straightforwardly captures the phenomenon of CC. Multiple climbing was accounted using the exact same reasoning, the difference being that more than one projects semantic information in the situation argument node in this case. The obligatoriness of CC in GSG is attributed to a general ban of clitics from appearing with infinitives. This auxiliary-like analysis of the specific two verbs in GSG is not available to any other of the verbs belonging to the class of restructuring verbs in GSG. Thus, all other verbs of the restructuring verbs will be parsed as regular complement taking verbs and as such will not induce climbing according to fact.

Acknowledgements

I'm indebted to my supervisor, Ruth Kempson, for invaluable comments and constructive critique on various parts of this paper. I also thank the audiences at MGDLT4 and LAGB 2009 for useful comments and discussion. Particularly I want to thank Ronnie Cann and Wilfried Meyer-Viol for helpful comments. The Arts and Humanities Research Council (AHRC), the Leventis Foundation and the Alexandros Miaris Board are gratefully thanked for providing partial funding related to parts of the work presented in this paper. Normal disclaimers apply.

References

- Blackburn P. & Meyer - Viol W. (1994). Linguistics, Logic and Finite Trees. *Bulletin of Interest Group of Pure and Applied Logics* 2, 2-39.
- Bouzouita M. (2008a). *The Diachronic Development of Spanish Clitic Placement*. Phd Thesis, King's College, London.
- Bouzouita M. (2008b). At the Syntax-Pragmatics Interface: Clitics in the History of Spanish. In: Cooper R. & Kempson R. (eds), *Language Evolution and Change*. Cambridge: Cambridge University Press.
- Cardinaletti A. & Shlonsky U. (2004). Clitic Positions and Restructuring in Italian. *Linguistic Inquiry* 35, 519-557.
- Cann R., Kempson R. & Marten L. (2005). *The Dynamics of Language*. Oxford: Elsevier.
- Cann R. (forthcoming). Towards an Account of the English Auxiliary System: Building Interpretations Incrementally. In: Kempson R. & Gregoromichelaki E. (eds), *The Dynamics of Lexical Interfaces*. CSLI publications.
- Chatzikyriakidis S. (2006). *Clitics in Modern Greek : A Dynamic Account*. MSc thesis, King's college London.
- Chatzikyriakidis S. (2009a). Clitics in Grecia Salentina Greek: A Dynamic Account. *Lingua* 119, 1939-1968.
- Chatzikyriakidis S. (2009b). *Clitic Climbing: A dynamic account*. Ms, King's College, London.
- Chatzikyriakidis S. In preparation. *Clitics in Four Dialects of Modern Greek : A dynamic account*. Phd thesis, King's college, London.
- Chatzikyriakidis S. (forthcoming). A Dynamic Account of Clitic Climbing: A First Sketch. In: Kempson R. & Gregoromichelaki E. (eds), *The Dynamics of Lexical Interfaces*. CSLI publications.
- Chomsky N. (1986). *Barriers*, MIT Press.
- Cinque G. (1999). *Adverbs and Functional heads: A Cross- Linguistic Perspective*. New York, Oxford University press.

- Cinque G. (2001). Restructuring and the order of aspectual and root modal heads. In: Cinque G. & Salvi G. (eds.), *Current studies in Italian Syntax*. Elsevier, 137-155.
- Cinque G. (2006). *Restructuring and Functional Heads*. Oxford Studies in Comparative Syntax, Oxford university press.
- Fodor J.A. (1975). *Language of Thought*. Cambridge: Harvard University Press.
- Katsoyannou M. (1995). *Le Parler Greco de Galliciano (Italie): Description d'une Langue en Voie de Disparition*. Phd thesis, University of Paris VII.
- Kayne, R., (1989). Null Subjects and Clitic Climbing. In: Jeggli O. & Safir, K.J (eds), *The Null Subject Parameter*. Dordrecht, pp. 239-261.
- Kempson R. & Cann R. (2007). *Production Pressures, Syntactic Change and the Emergence of Clitic Pronouns*. Ms., King's college London.
- Kempson R., Meyer-viol W. & Gabbay, D., (2001). *Dynamic Syntax: The Flow of Language Understanding*. Blackwell Publishing.
- Kempson R. & Chatzikyriakidis S. (2009). *The Person Case Constraint as a Treegrowth Restriction*. Ms, King's College, London.
- Manzini R. (1983). *Restructuring and Reanalysis*. Phd thesis, MIT.
- Miller P. & Sag I. (1997). French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15, 573-639.
- Monachesi, P., (1993). Object Clitics and Clitic Climbing in Italian HPSG Grammar. In: *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics*. Utrecht, 437-442.
- Monachesi P. (1998). Italian Restructuring Verbs: a Lexical Analysis. In: Hinrichs, E., Kathol, A., Nakazawa, T. (eds.), *Complex Predicates in Nonderivational Syntax*. *Syntax and Semantics* 30, Academic Press, San Diego, 313-368.
- Monachesi P. (1999). The Syntactic Structure of Romanian Auxiliary (and Modal) Verbs. In: Bouma G. et al. (eds), *Constraints and resources in natural language syntax and semantics*. CSLI publications, Stanford
- Ralli A. (2006). Syntactic and morphosyntactic phenomena in Modern Greek dialects: The state of the art. *Journal of Greek Linguistics* 7, 121-159.
- Rizzi L. (1982). Issues in Italian Syntax. *Studies in Generative grammar* 11. Foris Publications, Dordrecht.

A new sociolinguistic variable in Cypriot Greek

PANAYIOTIS A. PAPPAS

Simon Fraser University

papappas@sfu.ca

1. Introduction

The object of study in this brief report is the variable pronunciation of the syllable /liV/ in Cypriot Greek, i.e. in words such as /ðuliá/ ‘work’, /maliá/ ‘hair’, etc. As will be seen below, this type of variation has barely been noted in the literature, so it is perhaps appropriate to justify it as an object for variationist study. Tagliamonte (2006) uses the term *super token* to refer to instances in which a single speaker uses two different variants of a variable in the same utterance, because these exemplify the type of variation that is suitable for quantitative analysis. In this corpus, the best example of a super token can be seen in the following excerpt (1), where an educated female speaker in her mid-twenties, who is recounting a recent trip to Sweden, switches between a palatal lateral and a palatal fricative in the middle of a noun phrase:

- (1) iðame to-ðimarxio **tim-baja** **i-paʎa** i-poli
see.1PL.PAST DEF-cityhall.DO DEF-old.DO DEF-old.SUBJ DEF-city.SUBJ
ine gamlaston
be.3SG.PRES gamlaston
‘We saw City Hall, the old, “old city” is “gamlaston” (in Swedish)?’

The earliest mention of this variation comes from Christodoulou (1967), who locates this phenomenon in the cities of Lemesos and Amohostos.¹ More recently, Arvaniti (1999, to appear), also mentions this variation, but according to her it is more characteristic of the region of Larnaca. Newton (1972:24), however, does not mention this type of variation. According to him, [l] is a “voiced alveolar lateral, somewhat palatalized and long before /y/.” Thus, even though this variation has been noted in the literature, not much is known about it. I present here the results of a variationist analysis of the phenomenon, including a detailed description of the possible variants, two of which have not been identified up to this point. The results of the quantitative examination show that the innovative variants are favoured by male speakers and disfavoured by females, while a speaker’s level of education and place of residence are not important factors. In the discussion section, I make a preliminary attempt at explaining this pattern. I suggest that the innovative variants are a supra-regional feature of generalized Cypriot Greek that has covert prestige. At the same time, I highlight some aspects of the pattern that indicate that the meaning of the variable is more complicated, and suggest that further research employing the construct of *indexical field* could further our understanding. Overall, the emergence of this variable is seen as yet another sign of the rising status of Cypriot Greek within its native community.

¹I am grateful to Spiros Armosti for bringing this to my attention. This work has been supported in part by SSHRC Standard Grant 639510. I wish to thank the Department of English Studies at UCY for their hospitality during the Spring semester of 2007, the participant-interviewers, and all the participants. The usual disclaimers apply.

2. Methodology and Results

The database has been constructed on the basis of conversations, which were recorded during the spring of 2007 by eight participant-interviewers in Cyprus. These research assistants were students at the Department of English Studies at the University of Cyprus in Lefkosia, where I taught as a visiting professor during that period. Thus, I had the opportunity to train them in terms of fieldwork techniques, ethical standards etc. The team members were all native speakers of Cypriot from different areas of the southern part of the island, and were instructed to conduct conversations with family members and close friends.

The interviewers were not told the precise nature of the investigation; they were only told that I wanted to collect a database of vernacular Cypriot. Thus they were not prone to control their pronunciation or that of the participants, except for two instances at the beginning of a conversation where they instructed their mothers to speak Cypriot, instead of the standard.

Table 1. Age, Sex and Area of origin of participants

	12–17		18–35		40+		Totals
	M	F	M	F	M	F	
Lefkosia	1	2	3	10	1	0	17
Lemesos	0	0	3	10	0	5	18
Larnaka	0	1	3	4	3	3	14
Kokinohoria	0	2	0	0	0	1	3
Totals	1	5	9	24	4	9	52

The team was able to interview 44 participants, ranging in age from 12 to 80 years old. The interviews were recorded on a Marantz PDM 660, with a Sony ECB omnidirectional lapel microphone. For this analysis, I am able to use data from all 52 speakers, most of whom are female (38). There is a wide range of topics in the corpus, since the interviewers were given free range and actually instructed to follow the interests of their participants, in order to elicit as free-flowing a conversation as possible. Some of the more common topics are relationships between the sexes, football, politics, travel, and school. There are several clues that these conversations are casual in nature. There is much laughter and teasing, interruptions from other members of the family or calls on the cell-phone.

The classification of the participants can be seen in Table 1. The interviewers were able to recruit speakers in four different locations: Lefkosia, which is the capital and is located in the center of the island; Lemesos, the second largest city is located on the coast about 60km southwest of the capital; Larnaka, the third largest city which is also on the coast and 40km south of the capital; and the area of Kokinohoria, which is a collection of towns and villages 40km east of the capital. The participants can be divided into three peer groups (12-17, 18-35, 40 and older), but notice that there is only one male participant for the youngest group. Also as you can see not all ages are represented in every region, since there are only three speakers from the rural area of Kokinohoria, and two of them are teenagers. The best represented regions are Lefkosia and Lemesos.

Altogether, 966 tokens of (liV) were extracted from the database using Praat 5.1.2,

Table 2. The envelope of variation for (liV)

	Tokens	Coding
1. Geminate lateral [ʎ:]	142	0
2. Non-geminate lateral [ʎ]	397	0
3. Affricate [gʝ]	3	1
4. Geminate fricative [j:]	112	1
5. Non-geminate fricative [j]	293	1
6. Glide [j]	19	1

and were analyzed in terms of their pronunciation. According to Arvaniti (to appear), the variation is between the palatal lateral and a long (phonetically geminate) palatal fricative, which is realized as a glide only in weak positions, a finding that is partially confirmed in the present data. Arvaniti also notes that the variant lacks extensive voicing, which is also confirmed in this dataset. However, unlike Arvaniti, and Armosti et al. (2006)—whom she cites—the present data reveal a larger set of variants that may occur in conversational situations, including a non-geminate palatal fricative, and a palatal affricate as detailed in Table 2. Still, the lateral and fricative variants are the most frequently occurring ones. Since GoldVarb cannot execute multinomial analyses, the variants were grouped as 1 and 2 vs. 3, 4, 5, and 6; in other words the analysis was conducted along the contrast lateral (coded as 0) vs. non-lateral (coded as 1). Figures 1 and 2 depict the variants of the lateral and fricative pronunciations respectively. The short lateral in 1a is 0.054 seconds long, while the geminate in 1b is 0.178 seconds long. Similarly, the short fricative in 2a is 0.047 seconds long, while the geminate in 2b is 0.169 seconds long. In Figure 3 we see the other two non-lateral variants; in 3a, there is the characteristic burst of a plosive, while in 3b we see the sloping F2 of a palatal glide.

In addition to not having a truly balanced sample of speakers, another anomaly in the dataset is that the 966 tokens do not reflect many types. As can be seen in Table 3, most tokens are of /ðulia/ followed by /teliono/ (or one of its derivational or inflectional forms). Note also that [teʎ(:)a] or [tej(:)a] only represents the meaning ‘completely’ or ‘totally’. The form that means ‘perfectly’ is always pronounced [telia], i.e. with three syllables.

Table 3. Types and tokens in the dataset

Type	Translation	Tokens
/ðulia/	‘work’	184
/teliono/	‘finish’	145
/telia/	‘completely’	98
/malia/	‘hair’	65
/palio/	‘old’	103
/xilia/	‘thousand’	95
/yialia/	‘glasses’	54
Total		744

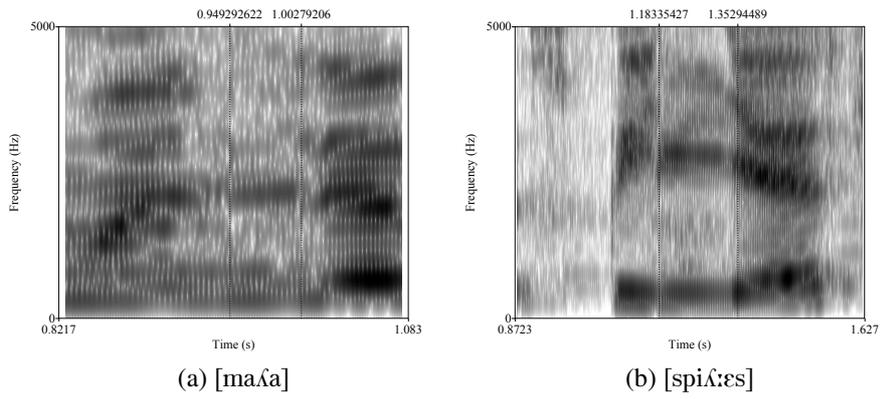


Figure 1. Spectrograms of lateral variants

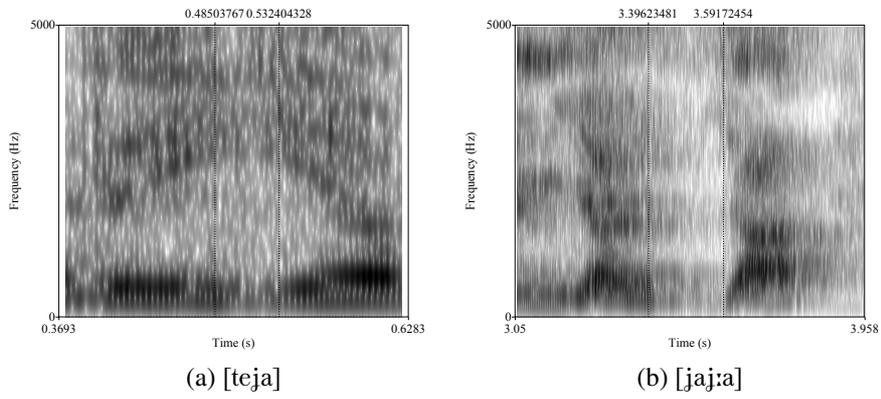


Figure 2. Spectrograms of fricative variants

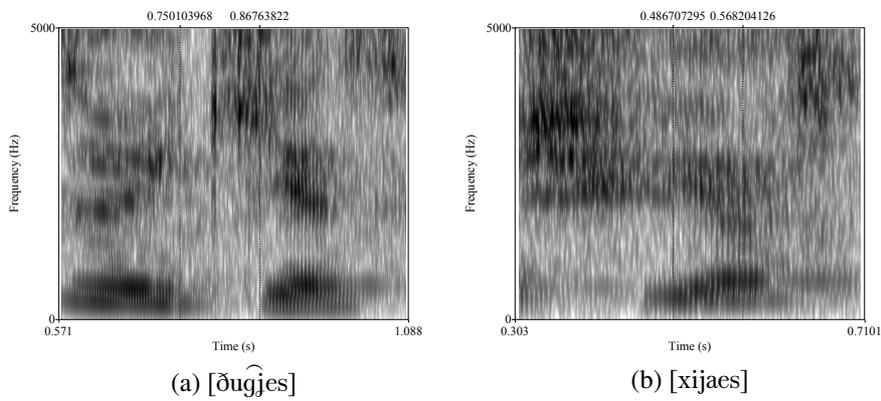


Figure 3. Spectrograms of affricate and glide variants

Despite being less than perfect when compared to variationist standards, the dataset does allow for a meaningful statistical analysis to be conducted. Let's begin by looking at the distribution of the innovative variant (the fricative) according to age. As Figure (4) shows, there is a dramatic increase in the use of this variant for speakers younger than 30.

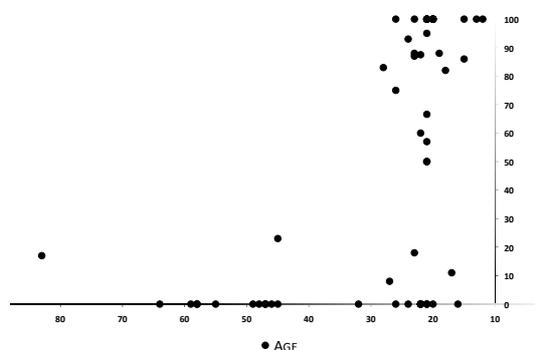


Figure 4. Usage (%) of the innovative (non-lateral) variant by age

The results indicate that we have a change in progress that began sometime in the past 10 to 15 years. Given that only speakers below the age of 30 participate in the change, the more detailed statistical analysis (GoldVarb) focuses on this group only. The following linguistic and non-linguistic factors were tested (Table 4).

Table 4. Set up of the variationist analysis

Linguistic	Non-linguistic
stress	sex
preceding vowel	education
following vowel	region
position in the word	years in Greece

Table (5) shows the results of the Goldvarb analysis. Numbers in angled brackets indicate tokens that had to be excluded. For the group *following vowel*, there were two tokens with /i/, both of which had a palatal fricative. For *preceding vowel* and *position*, there were three tokens with (liV) in word initial position. All three were lateral. Finally, in the group *region* the tokens from Kokinohoria speakers (50) were excluded. Based on the value of *Range*, the results show that in terms of linguistic factors the strongest group is the following vowel, with [a] and [e] favouring the innovative variant while the back vowels do not. Next we have preceding vowel, where [e] and [i] favour the innovative variant while [a] [o] and [u] do not. In terms of stress, [j(:)] is favoured when the syllable is stressed, while in terms of position, it is favoured in word-final position. I would like to emphasize, however, that these results come from a limited number of lexical items, so they may not hold up under more detailed investigation. In terms of non-linguistic factors, we see that speakers who have stayed for a length of time in Greece disfavour the innovative variant, as do female speakers. Male speakers on the other hand favour it. The

education of the speaker and the region that they live in do not have a significant effect. These results are discussed in more depth in the next section.

Table 5. Quantitative analysis results for the use of the innovative variant (non-lateral) of (liV), in Cypriot Greek.

Group	Factor	Weight	%	N	
Following Vowel*	a	0.57	58	507	
	<2>	e	0.57	70	67
		o	0.33	45	107
		u	0.06	5	25
<i>Range</i>		51			
Preceding Vowel**	e	0.71	60	223	
	<3>	i	0.59	56	81
		u	0.47	67	147
		a	0.33	48	282
		o	0.26	31	22
<i>Range</i>		45			
Stressed Syllable	stressed	0.57	58	489	
	unstressed	0.36	51	269	
<i>Range</i>		21			
Position***	final	0.54	58	574	
	<3>	medial	0.37	48	181
	<i>Range</i>		17		
<u>Non-linguistic factors</u>					
Years in Greece	no time	0.53	59	676	
	over a year	0.22	29	82	
<i>Range</i>		31			
Sex	male	0.66	67	175	
	female	0.44	52	583	
<i>Range</i>		22			
Education	advanced	[0.51]	55	560	
	basic	[0.45]	58	198	
Region****	Lefkosia	[0.51]	55	263	
	<50>	Lemesos	[0.49]	54	253
		Larnaka	[0.45]	53	192

3. Discussion

First, let us consider the type of sociolinguistic variable that (liV) represents. Within the accepted Labovian paradigm, there are three recognized types of sociolinguistic variables: indicators, markers and stereotypes. There is ample evidence in the conversations that the [ʎ(:)/j(:)] distinction is not treated as a stereotype. For example, in the following exchange (2), the interviewer hears something in her mother's speech that sounds like Standard Greek and so asks her to use Cypriot instead. In response, the mother repeats the phrase /me ta avya/ 'with eggs', pronouncing it [metafka] this time. Notice, however, that she does not change the pronunciation of [avɾɛʎa].

(2) θa tiyaniso avɾɛʎa metavya
 FUT fry.1SG.PRF asparagus.DO with.the.eggs

kipriaka se parakalo
 CypriotACC.SG you.DO please.1SG.PRES

metafka
 with.the.eggs

Mother: 'I will make asparagus with eggs'
 Daughter: 'Cypriot, please'
 Mother: 'with eggs'

On the other hand, it cannot be said that this variation is totally below the level of consciousness. Themistocleous (2008), in her dissertation on orthographic conventions for Cypriot online, mentions that <teleia> is spelled <teja>, reflecting a fricative pronunciation. Furthermore, in this dataset there is an instance of a fricative user switching to the palatal lateral in response to the interviewers use of the lateral at the beginning of the conversation in the word /ðulia/ (ex. 3).

(3) sti ðuja mu fonazun me k^hristi:n
 at work I.POSS call.PRES.3PL I.DO Christine

sti ðuʎa indalos se lalusi
 at work how you.DO call.3PL.PRES

sti ðuʎa lalun me k^hristi:n
 at work call.PRES.3PL I.DO Christine

Christos:² 'At my work they call me Christine (a joke)'
 Interviewer: 'What do they call you at work?'
 Christos: 'At work they call me Christine.'

Understanding the degree to which Cypriot speakers are aware of this pattern of variation is important, because this is the key difference between variables that are *indicators* and ones that are *markers*. Wolfram and Schilling-Estes (1998) state that because they operate at a different level of consciousness, markers are affected during style shifting whereas indicators are not. There are not many opportunities to study style shifting in the recorded conversations themselves. However, while wrapping up the project I conducted exit interviews with the student RAs. These are brief conversations approximately 15 minutes long, which are formal in nature. Because of the shorter length they do not include many tokens of the variable. Of the eight RAs four are users of the

²This and other speaker names are pseudonyms.

lateral variant and four are users of the fricative. There is only one possible case of style shifting. Interviewer Melpo, who uses the fricative consistently in the conversation with her mother, has three lateral tokens and only one fricative token during the exit interview, indicating, perhaps, that she is aware of the more formal situation and style shifts to match it.

There is then some limited evidence that speakers are aware of the variation, enough to style shift, making (liV) a sociolinguistic marker. It would be appropriate here to remind readers that Eckert (2008:463-464) cautions that the Labovian nomenclature may not always capture the full meaning of a sociolinguistic variable.

The difference between the notion of marker as used in variation studies and the index of Silverstein's treatment is in the ideological embedding of the process by which the link between form and meaning is made and remade . . .

. . . the reconstruals are 'always already immanent' ([Silverstein] 2003: 194) precisely because they take place within a fluid and ever-changing ideological field. The emergence of an *n* + 1st indexical value is the result of an ideological move, a sidestepping within an ideological field. In order to understand the meaning of variation in practice, we need to begin with this ideological field, as the continual reconstrual of the indexical value of a variable creates, in the end, an *indexical field*.

In a complex sociolinguistic environment, such as that of the Cypriot Greek speaking community, where the local variety (Cypriot) coexists not only with its diglossic counterpart (Standard Greek), but also with the language of its colonial past (English), as well as another official language, one that is politically and historically charged (Turkish), it is unlikely that the value of any sociolinguistic variable can be simply defined. Further research, focused on mapping out the indexical field of (liV) is required.

Finally we should consider how the results of this study fit in with what we already know about Cypriot Greek, and particularly what they signify about the current status of the variety. Terkourafi (2005) discusses the emergence of a modern koiné in Cyprus, which has taken hold throughout the island especially after the events of 1974. She reviews the pertinent literature and highlights 10 segmental features that identify this koiné, which I list below.

1. The palatalization of velar consonants before front segments.
2. The retention of long (geminate) consonants.
3. The aspiration of voiceless stops.
4. The retention and expansion of final /n/.
5. The devoicing of intervocalic and word-initial voiced stops.
6. The deletion of intervocalic voiced fricatives.
7. The epenthesis of /ɣ/ in verbs that end in /-evo/.
8. The dissimilation of obstruent + obstruent clusters into fricative + stop.
9. The change of /i/ to [ç] after /v/, /ð/, /θ/, /p/, or /f/.
10. The prothesis of /i/ for certain verbs.

The variable under discussion is not included in this list, and as I mentioned in the introduction, Arvaniti (to appear) considers the fricative pronunciation [j(:)] a dialectal (i.e. regional) feature, particularly associated with Larnaka. On the other hand, Christodoulou (1967) claims that it is a

feature of Amohostos and Lemesos. However, in this corpus, [j(:)] is robustly present in the three urban areas of Lefkosia, Lemesos and Larnaka as well as the rural Kokinochoria region, especially among teenaged speakers. On the basis of these findings, I would argue that the innovative pronunciation [j(:)] is not a feature of any local patois, but rather a supra-local feature, a feature of what Terkourafi calls generalized Cypriot Greek (gCG), the modern Cypriot koiné. Whether this feature has always been present in these areas or whether it has been spreading over the past two decades is a question for further research.

According to Terkourafi and others (cf. Terkourafi 2005:335), gCG is based on the regional lect of Mesaoria and has developed through simplification and leveling processes that are the result of contact between the various regional varieties. In addition to this, gCG is also characterized by a few innovative features such as the blended forms such as [xartca] instead of [xarca], ‘papers’ or the use of secondary stress with extrametrical clitics as in [to emvóliòn tu] instead of [to emvólión tu], ‘its vaccine’, leading Terkourafi to remark that

Such a wealth of new productive mechanisms and novel constructions is not what one expects of a retreating variety, and attests to the overall vitality of the Cypriot Greek dialectal continuum, though of course different elements may be falling out of use, as new ones emerge.

The data examined in this study reveal that the variant [j(:)] is an additional innovative feature of gCG. Moreover, it is independently motivated, a true native development within Cypriot Greek, whereas most of the innovative features mentioned by Terkourafi appear to have been influenced by contact with the standard variety. Considered from this perspective, some social aspects of the variation make sense. The results of the variationist analysis (cf. Table 5) showed that it is men who favour the innovative variant, whereas women do not. According to both Trudgil (1974) and Chambers (2003), features with *covert prestige* are more common among males than females as is the case in this study. As a local feature in a diglossic and overall charged sociolinguistic environment, it is more likely for [j(:)] to have covert rather than overt prestige. The covert prestige of this variant would explain the style shifting behaviour I discussed earlier, either as a response to lateral use by the interviewer, as in the case of Christos in example (3), or as a response to a more formal situation as in the case of the participant-interviewer who uses [j(:)] in conversations with friends and family but uses [ʌ(:)] during her exit interview. Finally, if [j(:)] is, indeed, a marker of covert prestige, Cypriot speakers, who have lived in Greece and have been educated there, may consciously avoid it because they are aware of its non-standard status.

On the other hand, the fact that a speaker’s education is not a factor in this pattern is unexpected in this scenario, because covert prestige variants are usually disfavoured by well educated speakers. The complicating factor here may be that the speakers for whom we are making this distinction are quite young in age (early 20s) and so their level of education may not be as important a predictor of linguistic behaviour as social network or personal identity considerations. Perhaps an investigation of the indexical values of this variable, in the spirit of Eckert (2008), would provide a clearer understanding of its meaning. Finally, I would like to suggest that the emergence of this variant can be seen as a linguistic indication of the rising status of gCG as this has been reported on the basis of metalinguistic evidence—cf. Papapavlou (1998), Arvaniti (2002), Tsiplakou (2004). The comparison between the earlier study and the more recent ones indicates that Cypriot speakers are becoming more self-confident, and are more positive towards their native variety. The development of a sociolinguistic marker that carries covert prestige could be interpreted as a sign that gCG is maturing into a robust vernacular and may yet become a standard, given the right political circumstances.

References

- Armosti, S., Christodoulou, K., Katsoyannou, M., and Protopapas, A. (2006), Cypriot Greek geminates: the contrast between [j] and [j:], Paper presented at the Colloquium of the British Association of Academic Phoneticians. Edinburgh, UK, April 10–12.
- Arvaniti, A. (1999), Cypriot Greek. *Journal of the International Phonetic Association*, 29, 2: 173–178.
- Arvaniti, A. (2002), Dimorfia, diglossia, ke i emfanisi tis Kipriakis Neoelinikis Kinis [Bimorphism, diglossia and the emergence of Cypriot Modern Greek Koiné]. In Clairis, C. (ed.), *Recherches en linguistique grecque: Actes du 5e colloque international de linguistique grecque (Sorbonne, 13–15 September 2001)*, Paris: L'Harmattan, 2002, pp. 75–78.
- Arvaniti, A. (to appear), A (brief) overview of the phonetics and phonology of Cypriot Greek. In *Proceedings of the conference on “The Greek language in Cyprus from antiquity to today,” University of Athens, 23-24 May 2008*, to appear.
- Chambers, J. K. (2003), *Sociolinguistic theory: linguistic variation and its social significance*. Oxford: Blackwell Pub, 2 edn.
- Christodoulou, M. N. (1967), Sintomos simiosis peri tis siniziseos en ti Kipriaki dialekto [A brief note on coalescence in the Cypriot Dialect]. *Kypriakai Spoudai*, 31, 93-97.
- Eckert, P. (2008), Variation and the Indexical Field. *Journal of Sociolinguistics*, 12, 4: 453–476.
- Newton, B. (1972), *Cypriot Greek: Its Phonology and Inflections*. Janua Linguarum, Series Practica, 121, Mouton de Gruyter.
- Papapavlou, A. (1998), Attitudes toward the Greek Cypriot Dialect: Sociocultural implications. *International Journal of the Sociolinguistics of Language*, 184, 1: 15–28.
- Tagliamonte, S. A. (2006), *Analysing sociolinguistic variation*. Cambridge and New York: Cambridge University Press.
- Terkourafi, M. (2005), Understanding the present through the past: processes of koineisation in Cyprus. *Diachronica*, 22, 2: 309–372.
- Themistocleous, C. (2008), *The use of Cypriot Greek in Synchronous Computer-mediated Communication*. Ph.D. thesis, University of Manchester.
- Trudgil, P. (1974), *The social differentiation of English in Norwich*. Cambridge [Eng.]: University Press.
- Tsiplakou, S. (2004), Stasis apenanti sti glossa ke glosiki alagi: mia amfidromi sxesi? [Stances towards language and language change: a mutual relationship?]. In *Proceedings of the 6th International Conference of Greek Linguistics, Rethymnon, 20 September 2003*, <http://www.philology.uoc.gr/conferences/6thICGL/ebook/g/tsiplakou.pdf>.
- Wolfram, W. and Schilling-Estes, N. (1998), *American English: Dialects and Variation*. Malden, MA: Blackwell.

The Reshaping of the Mediopassive Endings: Evidence from Modern Greek Varieties

NIKOLAOS PANTELIDIS
University of Athens
npantel@phil.uoa.gr

Abstract

The present paper discusses cases of morphological change in the paradigm of the mediopassive in Modern Greek varieties and the contribution of the observed changes to the theory of morphological change.

1. Introduction

1.1. The aim of the present paper is to present and analyze a number of changes which have led to the formation of the endings of the mediopassive voice in the Modern Greek dialects, as well as to connect them with linguistic theory in general and the theory of morphological change in particular. Needless to say, the discussion cannot cover neither the totality of the MG dialects nor all of the relevant changes.

1.2. The notion of inflectional paradigm will be considered basic for the issues discussed here; it is not within the aims of the present study to subject this notion to theoretical scrutiny. Every form (or “slot” or “cell” in some theoretical models) of the inflectional paradigm is connected to other forms via a semantic relationship (in broader or stricter terms: for example, with the sub-paradigm of the same number, with the same person in a different number or tense etc). The morphological relationship between the various forms of the inflectional paradigm is stronger or weaker depending on the phonological and/or semantic features they have in common. Most of the processes which cause changes on the form of the mediopassive endings follow this logic; the stronger the relationship between two forms, the higher the possibility of interaction between them. It is only in a few cases that the relationship between two interacting ‘cells’ of the paradigm is more distant, e.g. the same cell of the same tense but in a different person or voice¹. The present paper focuses mainly on the inflectional paradigm of the mediopassive imperfect of the traditionally termed “1st conjugation” (barytone) verbs, which presents greater variety of forms than the present.

1.3. The changes in the mediopassive endings of Modern Greek and especially its dialects can provide clues to the way in which speakers analyze their primary linguistic data, at least at the moment an innovation is created, and thus contribute to the better understanding of each “synchrony” (Booij 2007: 255, Joseph 2009: 53, 55)². At the same time, they can contribute to the development of the theoretical approaches to the mechanisms of morphological change, such as reanalysis and analogy, and show in which ways other morphosemantic and phonological factors may interact with each other, as will

¹ The to a great extent purely morphological distinction of voice does not seem to have diachronically impeded the possibility of interaction of forms belonging to the same lexeme, i.e. share the same lexical representation.

² “...diachrony is relevant to our understanding of synchronic systems. This is understandably so if one takes a ‘dynamic’ view of synchrony”.

be shown below. Finally, they contribute to the claim that the classic notion of morpheme is not sufficient to describe and interpret the changes in the mediopassive endings. In the present study, the traditional distinction between “dialect” (διάλεκτος) and “local variety / patois” (ιδίωμα) is not taken into account both for theoretical and for practical reasons.

1.4. The longer size (compared to those of the active voice) of the mediopassive endings renders them more liable to processes of restructuring, which may make them “producible” one from another or “relatable” to one another in the speaker’s mind and thus contributes to the creation of small-scale “generalizations” (see Joseph 2009). A case in point is the notoriously difficult reconstruction of the mediopassive endings of the Indo-European proto-language on the basis of the inflectional paradigms of the various IE descendant languages. These endings seem to be made up (or rather to have been made up) of smaller elements and are restructured in the various IE languages through complex analogical/ morphological processes (cf. Clackson 2007: 142-151 among others). What is more, in the case of the Modern Greek dialects in particular, the absence, at least in previous periods, of a linguistic standard has facilitated changes. The mediopassive endings thus present a considerable amount of local variation, although the various inflectional paradigms could be subsumed under a small number of basic types according to criteria such as stress pattern, following e.g. the practice of Newton’s 1972 article. Even within the system of each individual local variety is it possible in some cases to observe extensive allomorphy, not only phonological or morphological (Ralli 2005:67), but also connected to various factors as well as to the general history of the language (and occasionally to the influence of Standard Modern Greek). The so-called free variation of allomorphs of the mediopassive endings deserves a more in-depth study from the viewpoint of whether it may be influenced by factors such as style or phrasal/sentence rhythm³. Inflectional paradigms like those presented in prescriptive grammars of Standard Modern Greek are rare. The impression of inflectional paradigms without variation and allomorphy is sometimes created indirectly through publications (specialized or not) which provide partial or complete description of dialects (see also Newton 1972):

Mediopassive imperfect (singular) in the variety of Kea (Kollia 1933: 278)

(Table 1)

Island of Kea (Tzia):

ímudan ~ ímane ~ ímuna

ísudan ~ ísane ~ ísuna

ídane ~ ítane ~ ítone

Mediopassive imperfect in the dialect of Patmos (Papadopoulou 2005: 177-8):

(Table 2)

³ For example, the choice of the 3.PLUR. allomorph ‘-o(n)dan or -ó(n)dusan.

Island of Patmos:

3.SG: -útane ~ -údone ~ '-udane ~ -ódane

1.PL: -úmeste ~ -úmastóne

1.5. Newton in his 1972 article gave a first, quite detailed overview of the dialectal and geographical distribution of the various types of inflectional paradigms, as well as of the processes which led to the attested forms, based of course on a limited, by today's standards, amount of data. The evidence he gathered from his informants gives a different picture from the one deriving from the examination of the extensive material available in the contemporary data collections and corpora; this is due both to the imprecise and occasionally outright wrong answers of his informants, and to the extensive allomorphy present even within the same settlement (as pointed out above), something which does not come through in the data he sets out. Moreover, in my view, the distinction between diachronic processes and synchronic rules of the generative model of the period is less than clear in this paper. Parts of the data presented requires revision, while all the changes are viewed as simple changes on the level of form (as some of them undoubtedly are) without reference to the marking of grammatical categories. Yet the latter point cannot be ignored, and seems to constitute a crucial factor in several cases (cf. Janda & Joseph 1992, Joseph 2005, 2006, 2008, 2009). Speakers seem in some cases interested to emphasize specific grammatical categories which they perceive as insufficiently marked in some forms of the inflectional paradigm or whose morphological expression has become opaque, although in other cases this need is not evident. Of course one should be careful not to confuse the trigger of a change with the result it has on the inflectional paradigm; in any case, however, it is necessary for any interpretative attempt to take into consideration the whole inflectional paradigm as well as data from language history. The a-historical perspective on linguistic questions such as the one under discussion can lead to questionable conclusions which may also have repercussions on the theory, as will be shown in the case of the 3.PL. ending '-ondan of the mediopassive imperfect. The various linguistic varieties constitute, as Newton indirectly concluded, different stages of evolution of the inflectional paradigm in different "branches". Furthermore, the investigation of the structure of the inflectional paradigm in each dialect can provide important clues for the relative chronology of the changes.

2. The morphological structure of the mediopassive present and imperfect in Standard Modern Greek.

The forms of the inflectional paradigm of the mediopassive (present and) imperfect in Standard Modern Greek are structured as follows:

BASE + INFLECTIONAL SUFFIX. The inflectional suffix can be viewed as also containing the element to the left of the agreement markers (person+number) which displays in the present an alternation /o/ (or /u/) ~ /e/, while appearing in the imperfect as a columnally stressed /o/ (except for the 3.PL. allomorph '-ondan), which derives from the Ancient Greek so-called "thematic vowel":

(Table 3)

PRESENT	IMPERFECT
'-ome	-ómun(a)
'-ese	-ósun(a)
'-ete	-ótan(e)
-ómaste	-ómaste -ómastan
'-este -ósaste	-ósaste -ósastan
'-o(n)de	'-o(n)dan -ó(n)dusan

The precise analysis of the “thematic vowel” is a disputed issue due to the different approaches adopted depending on the model of morphological analysis (in general or of the Modern Greek verbal system in particular)⁴. As far as the dialects are concerned, the system of each should be examined separately in order to isolate the factors which determine its appearance, form and function. For example, while in Standard Modern Greek the thematic vowel appears regularly in the mediopassive imperfect as carrier of the stress (with the exception of the 3.PL. allomorph '-*ondan*), in the dialects placement of stress varies, and so does the form of the thematic vowel, which alternates between /o/ or /u/⁵ and /e/ (or /i/ in northern varieties) as in the present. The form of the thematic vowel in the present is determined in Standard Modern Greek by the combination of the morphosyntactic properties of person+number, while in the imperfect its form is stable, a result of gradual changes which have not yet been completed in all Modern Greek linguistic varieties.⁶ The basic stress pattern of the present, inherited in general from previous phases of the history of the language, requires recessive stress (on the antepenult) in Standard Modern Greek. In the imperfect, it requires stable stress on the thematic vowel. The inflectional suffixes can be considered “portmanteau morphs” in that they are carriers of the morphosyntactic properties of person+tense+number, perhaps even of verbal aspect since in the mediopassive aorist (i.e. the perfective past) the inflectional suffixes are identical to those of the active voice. The contrast ‘present : past’ is neutralized in the allomorphs *-maste* and *-saste* of the 1-2 PL.

3. The original inflectional paradigm of the mediopassive imperfect

The inflectional paradigm of the mediopassive imperfect which can be considered as lying at the origin of the inflectional paradigms of the different MG varieties, and which is recoverable not only through the direct sources of past forms of Greek (including older dialect sources) but also through the comparative study of the dialects in the framework of the historical-comparative study of genetically related language-forms in general, is as follows:

(Table 4)

-ómin
'-eso → '-eso(n)
'-eto → '-eto(n) (and '-oton)
-ómeθa → -ómesta(n) (→ -ómesθen/-ómeste(n) etc.)
'-este

⁴ See Ralli 1987: 258, Mackridge 1990: 269-277. A discussion on the precise morphological status of the thematic vowel in the Modern Greek verb is beyond the scope of the present paper.

⁵ /u/ even in southern varieties. Its presence there cannot of course be attributed to the law of raising of unstressed /o/ as in northern varieties.

⁶ In some approaches to the structure of the MG verb, it is considered in the imperfect as one of the exponents of “past”.

‘-ondo → *‘-ondo(n)* (*‘-ondan, -ónd-asi(n)* → *-óndisan*⁷)

Of course it cannot be assumed that the paradigm was completely homogeneous in the whole geographical area in which Greek was spoken, in fact the detailed historical-comparative investigation could even uncover larger geographical sub-sets of inflectional paradigms within (early) Medieval Greek (cf. Horrocks 2010:320-323).

4. General observations on the form of the 3.SG.

The retention in many dialects (or in their older sources) of the specific inflectional suffix, either unaltered or with changes in its vocalism, but always maintaining its original stress, is in line with its observed high frequency and therefore the assumed autonomy and strength of the 3.SG (in the sense of Bybee 1985) as a model of analogical change. In many varieties the stress pattern of the 3.SG was extended to the 1.SG which was originally stressed on the “thematic vowel” (perhaps under the combined influence of the more basic present, e.g. 1.SG *-ómin* → *-ómun*, but also *‘-umun*, compare 3.SG.IMPERF *‘-eton*, 1.SG.PRES. *‘-ome* / *‘-ume*), while in others (e.g. many northern dialects) the 3.SG underwent syncretism with the 3.PL. in *‘-undan*.

In fact, in older sources (of the 16th-17th c.) of dialectal varieties which nowadays present columnal accent on the “thematic vowel” as in Standard MG (e.g. in the Heptanese and in the Peloponnese: *-ómuna/-e, -ósuna/-e, -ótuna/-ótane*) it is still possible to find 3.SG forms stressed according to the “older” pattern, e.g. *‘-oton* / *‘-otun*. In parts of the Peloponnese and Central Greece one even finds *‘-etan* (and with northern vocalism: *‘-itan*), e.g. *estéketan káθitan érxitan* etc. In various today’s insular varieties (e.g. in the Dodecanese, Ikaria, Crete etc.) one finds in the imperfect forms ending in *‘-umu(n(e))* (1.SG), *‘-usu(n(e))* (2.SG), *‘-eto(n)/ ‘-edone/ ‘-uta* (3.SG) etc., and in various northern varieties forms in *‘-uman ‘-usan ‘-undan*.

While in some of these varieties the 1.PL (and the 2.PL wherever we have extension to *-ósaste* from original *‘-este*) has maintained stress on the “thematic vowel”, i.e. *-ómastan -úmaston* etc. (and with stress shift *-omástane -omástene* etc. as required by the trisyllabic window), in other varieties (e.g. northern ones) a tendency for fixing stress on the verbal base can be observed, with development of secondary stress due to the trisyllabic window, e.g. *káθumástan káθómasténe*, or with vowel deletion due to the same rule, e.g. *káθ‘masthan* or *káθum‘stan* etc. The fact remains that the 3.SG played a crucial role in the general development of these varieties. Of course it too underwent changes triggered by other forms of the inflectional paradigm, mainly by the equally strong 3.PL (see below):

a) Syncretism of the 3.SG and the 3.PL in the direction of the latter in many northern varieties among others (see also Ruge 1973:154-157)⁸. This is perhaps connected with the retention of the original stress pattern, which is identical to that of the 3.PL. in *‘-ondo*.

b) Extension of *-an* from the 3.PL. in *‘-ondan* to the 3.SG. and creation of *‘-etan* (e.g. in Peloponnese) or *-étane/-étani/‘-itan* in Old Athenian and in other varieties of Central Greece –“Sterea Ellada”- and Euboea). Interestingly, no or very little influence of the 3.SG on the formation of the 3.PL. is observable in the material examined.

c) Extension of the vowel /u/ from the 1.-2.SG *-ómuna -ósuna* to the 3.SG, yielding *-ótuna* in Heptanesian. As mentioned above, however, in older phases at least the original stress pattern was maintained (*‘-oton* or the even more archaic *‘-eto(n)*). Thus the 3.SG showed remarkable resistance before the modern form *-ótan(e)*, whose vocalism and stress pattern are unconnected to the original ending, finally prevailed.

The varieties which belong to each type are not necessarily genetically related (or at least, this characteristic does not constitute sufficient explanation on its own), while many of

⁷ See Pantelidis (2005).

⁸ In today’s Standard Modern Greek usage of some parts of northern Greece one may also observe the reverse direction of syncretism, i.e. the use of forms in *-ótan(e)* with plural function, e.g. *aftés erxótane* ‘they (FEM.) were coming’.

them do not represent a pure type; this shows the diachronic fluidity of the classification in the one or the other type as well as the constant appearance of new tendencies for restructuring of the whole paradigm or parts thereof. Thus, in some of the varieties which historically represent an inflectional paradigm of this type, the extension of forms with final /e/ has led to shift of stress, as required by the trisyllabic rule: *-úmune -úsune -údone* etc.

5. The form of the 3.PL: Its genesis and its role in further changes.

5.1. The 3.PL seems to have constituted an equally powerful analogical model, which lies at the origin of the creation and spread of the pattern 'present -e: imperfect -an'; this pattern spread to the 3.SG and in several northern and other varieties led to 1.-2.SG forms in *-man -san*:

(Table 5)

Northern varieties:

- 1.SG *-óman*
'-uman
- 2.SG *'-esan > '-isan*
-ósan
'-usan

Southern varieties (Euboean, Old Athenian, Megarian):

- 1.SG *-ómane*
- 2.SG *-ésane*
-ósane

According to Babiniotis (1972:204-206), the genesis of the 3.PL in *'-ondan* can be viewed as part of the general tendency towards "unification of the past" at the level of endings, but, as he himself admits, the expected result would rather be forms with a vowel /a/ marking 'past'⁹ immediately preceding the inflectional suffix (**-a-maste, *-a-(sa)ste, *-a-ndan* or rather **-a-nde*), following the pattern of the active voice (e.g. 1.PL. *-a-me*, 3.PL. *-a-n(e)*) although a "correct" linear ordering does not seem to be always the aim of the speakers (Joseph 2008:3):

(Table 6)

- [-past] *-u-n(e) : -o/u-nde*
- [+past] *-a-n(e) : *-a-nde (ή *-a-ndan?)*

A change along these lines seems to have taken place in Grico (Puglia, S. Italy, see Karanastasis 1982: 84), where /a/ as a marker of the past was transferred to the mediopassive imperfect in a position immediately preceding the suffixes denoting person + number (*'-a-mo, '-a-so, '-a-to, '-a-mósto, '-a-sósto, '-a-tto*). These forms could of course also be analyzed as signaling past also through the vowel /o/ (contrasting with /e/ which appears at the right edge of the present forms, e.g. *-ome -ese* etc.).¹⁰

5.2. If one insists in interpreting the genesis of *-ondan* as a replacement of the *-o* of original *'-ondo* through the marker *-an* of the 3.PL, which belongs to a set of markers of person+number which are unmarked for voice (sometimes more carefully reference to the influence of *-an* in the change of *'-ondo* to *'-ondan* is made, see Joseph 2006:2), then this change should be classified as a case of affix pleonasm (*-ond- + -an*), since at least the categories "person" and "number" (but tense as well), are marked on both elements participating to the creation of *'-ondan* (Joseph 2005). However, it is hard to see the original 3.PL. form as morphosemantically opaque, since tense (together with person and number) is sufficiently marked (tense also through the contrast /e/:/o/, e.g. *érxonđe : í-*

⁹ Babiniotis characterizes /a/ as the 'thematic vowel of the past' (1972:207-208).

¹⁰ An instantiation of what has been termed *extended exponence* (cf. Booij 2007:116,313, Coates 2000:622-623).

/érxondɔ), and therefore the conditions which according to Haspelmath (1993:297-298) lead to pleonasm do not seem to apply. According to the same author, “pleonastic affixation” consists in the addition of a productive affix onto a word in order to achieve more transparent marking of the morphosyntactic category which is already expressed in this word through a different morpheme, which however has become opaque¹¹. It is therefore a mechanism increasing the morphosemantic transparency of synchronically “irregular” or “unproductive” structures. The question in this case is in what sense a form like *í-/érxondɔ* could be considered irregular or unproductive at the time of its change to *érxondan*. Lehmann (2005: 141), providing a more sound perspective, speaks of hypercharacterized forms, which are created due to paradigmatic pressure:

“All of these examples [of hypercharacterization in inflection] clearly involve analogical transfer of a marker from a context in which it is the only operator to fulfill the function in question to a context where it pleonastically duplicates an operator already applied. We may generalize that hypercharacterization in morphology itself is based on analogy. Moreover, in a diachronic perspective, the two concurrent markers are not on the same level. There is an inner marker *which for some reason does not quite do the job*, and an outer marker which is currently productive and which speakers feel should appropriately appear on such a word form. A more precise formulation of the analogical account might therefore say that hypercharacterization is a kind of adaptation of a stem or word form based on paradigmatic pressure”.

In fact, according to him (2005: 151, fn. 22), “an analogical model does not need to be perfect in motivating each and every feature of the transformed item; it suffices that it share some features with the latter”. In the case under discussion, the active voice, which can be viewed as the semantically unmarked member of the system of voices in Modern Greek, must have provided the model, despite the fact that the contrast in the 3.PL. in the active voice is [-past] *-un(e)* : [+past] *-an(e)* (and/or *-asi*) (see above). The main question in this context remains why the “inner marker” “does not quite do the job”.

We are dealing here with a classic example of how the lack of attention to the historical record and to the dialectal data as collected up to the 20th c. can lead to erroneous conclusions. In older sources (of the 16th and 17th centuries)¹² one finds 3.PL imperfect forms in *-ondon* (beside the more recent *-ondan*). This ending, which is probably preserved in mainly insular varieties as *-o/u-don(e)*, came about as follows: The strong analogical model of the 1.SG. *-ómin* > *-ómun* (according to Horrocks 2010:321 also of the 3.SG. in *-en* bearing the so-called *ny ephelkystikon*) which has final /n/ influenced the nearest slots of the paradigm on a purely formal level, giving 2.SG : *-eso* → *-eson*, 3.SG *-eto*

¹¹ Haspelmath’s description of the phenomenon involves a contradiction, in that when a marker has become opaque, its recognition concerns mainly past synchronies and not the time of the appearance of the innovation. In other words, at the time when such an innovation is created, it is doubtful whether the speaker at least can be considered as capable of synchronically recognizing a marker which transparently expresses a category. In many cases, as e.g. in the change of the Latin infinitive *esse* (etymologically *es-se*) to *esse-re* in Vulgar Latin (cf. Ital. *essere*, Fr. *être*, Span. *ser*), it is in my opinion doubtful whether there still exists synchronically any marker of the infinitive, opaque or not (cf. Haspelmath 1993:299).

¹² E.g., the sermon of Maximos Peloponnisios, Ioannikios Kartanos etc. In the text of the Chronicle of Morea, as transmitted by the Copenhagen manuscript, as well as in the War of Troy, only *-ondan* (and *-óndisan*) is attested, which shows the chronological priority of the genesis of *-ondan* with respect to the 3.SG. *-tan* (from older *-ton*). Editions of the texts: (a) Nikolopoulou A. (1995). “Μαξίμου του Πελοποννησίου εξήγηση του ‘Κατά Ιουδαίων’ έργου του Μελετίου Πηγά”. Parnassos 37:308-346. (b) Ιωαννίκιος Καρτάνος, Παλαιά τε και Νέα Διαθήκη (Βενετία 1536). Ed. by Eleni Kakoulidi-Panou & Eleni Karantzola. Thessaloniki: Kentro Ellinikis Glossas, 2000. (c) Το Χρονικό του Μορέως, ed. by Petros Kalonaros. Athens: Ekati editions. (d) Ο πόλεμος της Τρωάδος, ed. by M.Papathomopoulos & E.M.Jeffreys. Athens: Morfotiko Idryma Ethnikis Trapezis, 1996.

→ *-eton*, as well as 3.PL. *-ondo* → *-ondon*. *-ondon*, which shared the feature [+past] and the final /n/ with the unmarked for voice marker of person + number (3.PL) *-an*, was further transformed into *-ondan* under the influence of *-an*. Of course a more exhaustive investigation of older (late Medieval/ early Modern) Greek texts and dialectal varieties could lead to more reliable and detailed answers.

5.3. In the opposition *-onde* : *-ondan*, the elements *-e* and *-an* were reanalyzed as markers of the category of tense ([+past]), since the presence in both forms of the element *-nd-* could mark the categories of person and number (see also Mackridge 1990:276).

The stage which included 3.SG *-eton* and 3.PL *-ondon* probably caused the appearance of the 1.-2. PL. markers *-maston(e)* *-saston(e)* *-meston* etc. (with thematic vowel /ó/ or /ú/ or /u/) which are recorded in various (mainly insular) varieties (Patmos, Kythnos, Symi, Crete, Rhodes¹³, parts of the Peloponnese etc.), and which must be due to a similar process of reanalysis at this precise stage. Furthermore, the 1.-2. PL. forms in use in many parts of the Peloponnese, *kaθómastun*, *kaθósastun* (similar forms are attested also in Velvendo, prefecture of Kozani –Macedonia-, e.g. *érxumástun érxusástun*) derive from corresponding processes of reanalysis on the basis of the contrast *-e* : *-un* in the 1-2.SG.:

(Table 7)

íme : *ímun* → *ímaste* : *ímastun*
íse : *ísun* → *ísaste* : *ísastun*

Patterns which are the result of an initial change triggered by different causes are reinterpreted by the speakers, often without taking into consideration the overall morphological structure of the language at least as would be analysed by linguists:

Stage 1. *-ondo* → *-ondon* → *-ondan* (3.PL.IMPERF)

Stage 2. PRES *-onde* : IMPERF *-ondan* → reanalysis of *-e* and *-an* as markers of tense: PRES *-e* : PAST *-an* (*-nd-* : 3.PL mediopassive)

Stage 3. Extension of the pattern to the 3.SG. which shares the feature of person with the 3.PL : PRES *-te* : IMPERF *-tan* (← *-to(n)*).

Stage 4. In northern (and other) varieties, extension to other forms of the paradigm:

1.SG *-me* : *-man*
 2.SG *-se* : *-san*

In fact, from the moment that the innovative forms come into existence on the basis of their analogical models, the former can gradually acquire autonomy with respect to the latter as to several parameters, e.g. addition of the vowel /a/ in 1.2.SG but of the vowel /e/ in the 1.2.PL, e.g. π.χ. *í-mastuné isastuné* ('we were, you were') versus *í-muná í-suná* ('I was, you (SG) were').

5.4. In such cases, the notion of morpheme is not sufficient for the interpretation of the developments. Janda and Joseph 1992 (cf. also Joseph 2009:52-53), discussing the recurrent partial similarities between various forms of a paradigm do not accept hyper-segmentation into morphemes, which would go against the elsewhere condition, but instead recognize elements (which cannot fit into the classic notion of morpheme) introduced on the basis of "meta-redundancy" rules, while several other scholars accept the existence of sub-morphemic units as necessary for morphological analysis¹⁴, both in inflectional and in derivational morphology. A slightly different approach is adopted by Bybee (1985:127-129), who views morphological structure within the framework of connections between lexical units or between forms of the inflectional paradigm, which do not function on the basis of a strict segmentation into morphemes. In this framework however one may recognize elements which would not be considered morphemes *stricto sensu*, but can nevertheless be viewed as markers of grammatical categories. Psycholinguistic (and neurolinguistic) research also provides interesting insights into the

¹³ Cf. e.g. Newton 1972:281. 3.PL. *érkund-e* : *érkund-on* → 1-2. PL. *érkumest-on érkust-on* (Rhodes).

¹⁴ See Luschützky 2000, Kubrjakova 2000 with extensive overview of the issue and bibliography.
 e-Proceedings of 4th Modern Greek Dialects and Linguistic Theory

way speakers process the structure of words and consequently into the basis of morphological change: Experiments have offered important corroborative evidence for the largely emergent character of morphological structure and for what speakers actually treat as meaning/function bearing units. Their processing does not necessarily conform to the morphological analyses linguists would come up with (see among others Devlin et al. 2004, Taft & Kougioussis 2004).

An interesting case is also provided by Ruge (1973:131, fn. 15), operating in much the same way: in this case as well, parts of the inflectional forms, which would in all probability not be considered as markers under a classic morphemic analysis (in this case /o/ vs. /e/) are perceived by the speakers as function-bearing units:

“The 2.PL. form [e]sterísθo (or [e]sterísto)), occasionally heard in place of *esteríte* is strange. I interpret it as an analogical formation on the basis of 3.SG *esterító*:

	3.SG.	2.PL.
PRES. (-e)	steríte	steríte
IMPERF. (-o)	(e)sterító	(e)sterísto”

I have recorded both *esterísto* as well as *aníxesto* (i.e. *anexósastan* ‘you were tolerating’).

6. 1.-2. SG structures of the type *-mu-tan(e)/-tone -su-tan(e)/-tone etc.*

In other changes, morphosemantic transparency seems indeed to increase, according to the principle one meaning : one form (as far as possible). Thus in many varieties, both northern and southern, the reanalysis of the inflectional suffixes of the 1. and 2.PL as containing the oblique weak inflectional forms *mas* and *sas* of the personal pronouns (see Ruge 1984) has led to the creation of 1.-2. SG forms like *ekimúmutóne* (standard *kimómun* ‘I was sleeping’), *ímutáne* (standard *ímun* ‘I was’), *ísudan* (standard *ísun*, ‘you (SG.) were’), *kaθóm’dan* (standard *kaθómun* ‘I was sitting’), *kaθós’tan* (standard *kaθósun*) etc. (see also Pantelidis 2006:290-292) analyzable as ‘thematic vowel+marker of person/number + *-tane*’: The creation of these forms seems to confirm Ruge’s theory much more than the changes in Standard Modern Greek. The problem in this context is whether after the reanalysis one is dealing with a sequence of two markers (e.g. *-mas-tane*, *-m(u)-tane*) or just one. The first solution, although it conforms to the speakers’ analysis of the 1.-2.PL. forms runs up against the difficulty of attributing a specific and clear function to the second element (*-tane* = [-per-fective], [+past], [-active] or combination thereof? ‘empty morph’?), something which is not always possible (Bybee 1985:128, Luschützky 2000:456-458, Kubrjakova 2000:424-425). It would also run counter to an important feature of Modern Greek verbal morphology, according to which the agreement properties which are important for syntax, i.e. person + number, are expressed on the right edge of the verbal inflectional form. On the other hand, it is obvious from the reanalysis that the speakers have isolated *-mas-* and *-sas-* in the forms of the 1.-2.PL. as markers of person+number. This is yet another case where the classic notion of morpheme cannot describe morphological structure adequately, since the new parsing made by the speakers identifies new markers of person+number on the basis of form and meaning similarities with elements outside the verbal inflectional paradigm. Moreover, it “disregards” both the overall morphological system of Modern Greek, which requires final position of the agreement markers in inflectional forms, and the syntactic congruity of such an analysis, since the pronominal forms *mas* and *sas* which were analyzed as bearers of meaning in the mediopassive forms do not represent the case of the subject, being genitive-accusative and not nominative forms. Interestingly, the dialectal varieties in which these forms are attested do not seem to employ 3.PL forms in *-óndusan*. This fact weakens the possibility that the latter was created on the basis of analyses of the 1.-2. PL. as *-mas-tan -sas-tan* → *-ón-tus-an* (*tus*: oblique form of the 3.PL personal pronoun), as has been suggested (see

Joseph 2008, 2009), while at the same time exemplifying the autonomy (in the sense of Bybee 1985) of the 3rd person forms and their consequent resistance to restructuring.

7. “Affix pleonasm”.

In this section cases are discussed which would fit into the concept of “pleonastic affixation” as conceptualized by Haspelmath (1993). The results are “hypercharacterized” forms in the sense of Lehmann (2005).

7.1. In the dialectal variety of the village Ochthonia in Euboea, the following inflectional paradigm of the mediopassive imperfect is attested (Favis 1911: 58):

(Table 8)

erx-úm' tane < **erx-úmutane* ‘I was coming’ (see above for similar structures)
erx-és' tanes (from older **erx-és' tane* < **erx-ésutane*)
erx-é tane
erx-úmastane
erx-ésastane
erx-ú(n)dane

In the 1.-2.SG, which came about through the process described in the preceding section, the deletion of /u/ in the otherwise southern dialect has “corrected” the violation of the trisyllabic window, leading to the 2.SG form **erx-és'tane*. The new form is clearly distinguished from the 1.SG through the form of the thematic vowel and /m/, as well as from the PL forms through the increased phonological difference. But the distinction from the 3.SG is not clear-cut: The increased phonological similarity to the 3.SG *erxétane* due to the form of the thematic vowel, the position of stress, the deletion of /u/ and the presence of *-tane* rendered the form opaque as to the category of person within the singular. This creates the conditions necessary for “affix pleonasm” as described by Haspelmath and Lehmann (see § 5.2 above); the addition of final *-s*, a marker of 2.SG unmarked for voice, restored the transparency of the form **erxés'tane*.

7.2. In some Euboean varieties 3.PL present forms in *-ondes/-undes* are attested, e.g. *léu(n)des pa(n)drévyo(n)des* (standard *léyo(n)de pa(n)drévo(n)de* ‘they are named, they get married’), which Minas (1987:47) indirectly but correctly, in my view, attributes to an older **-o(n)de-si*. The creation of the latter is probably quite old, belonging to a period when both the active and the passive voice displayed alternation between allomorphs ending in the element *-si* and allomorphs without *-si*:

(Table 9)

-un ~ -usi *-o(n)de ~ *-o(n)de-si > -o(n)des*
-an ~ -asi *-o(n)dan ~ *-ó(n)d-asi(n) → -óndisan / -ó(n)disáne¹⁵*

This case would be a more characteristic instance of what Lehmann terms “hypercharacterized forms” (2005:141), at least in the initial phase before the deletion of final /i/, in that a new marker *-si* of the 3.PL, unmarked for voice, was added onto the already extant marker of this category.

7.2. Another possible case of affix pleonasm is constituted by the mediopassive imperfect inflectional paradigm of parts of Aetolia, as reported by Papadopoulos (1927:93):

(Table 10)

-um'n-an
-is'n-an
-itan
-umast-an
-i(sa)st-an
-und-an

¹⁵ Cf. also Pantelidis 2005.

The *-an* contrasting with *-e* (e.g. *-und-e* : *-und-an*) in the 3.PL was reinterpreted as marking tense ([+past]) and was extended to the whole paradigm of the imperfect. Its extension to all the forms of the inflectional paradigm, in combination with the fact that in the singular it seems to have been added onto the whole original marker of person+number (e.g. in *-m'n-an* *-s'n-an*, where *-m'n-* < *-mun* και *-s'n-* < *-sun* through high vowel deletion, in contrast to what happens in other northern varieties, e.g. *-óman/'-u-man* and *-ósan / '-usan*) and did not replace the final phonemes of the older markers could be viewed as a reinforcement of the markers of the 1.-2.SG. These must have been at some point rendered partially opaque, or phonotactically unacceptable through the operation of phonological processes such as high vowel deletion ((?)**-m'n* < *-mun* and (?)**-s'n-* < *-sun*), if of course what lies at the origin of *-m'n-an* *-s'n-an* is indeed **-m'n* and **-s'n* and not *-m'na* *-s'na* (< *-muna* *-suna*). In the latter case, the interpretation of these forms should be different.

7.3. G. Salvanos (1918:14, fn.1) mentions a case from the variety of Corfu. According to him, many speakers in the city of Corfu employ 1.PL. mediopassive forms in *-omáste-me* (instead of *-ómaste/-omástene*). We are in all probability dealing (if the ending was correctly recorded) with the same mechanism here, which leads to a unified marking (in this case of the 1.PL.) at the right edge of the inflectional form, despite the fact that the categories person+number are already marked by the inflectional suffix *-omástene*. This development was perhaps facilitated by the phonological similarity of the syllable *-ne* (which probably evolved partially through phonetic processes, i.e. the addition of final /e/ onto the older ending *-mesten/-masten* due to the well-known tendency for open final syllables) with the unmarked for voice inflectional suffix of the 1.PL. *-me*.

The cases under discussion in this section present similarities with the process that Booij (2007:273-275) termed *systematization* (which leads to “overcharacterization”), referring to processes of derivation and not inflection. Koefoed & van Marle (2004:1581) view such processes as a type of *morphological adaptation* operating on the “output” and not “on the rule system as such”. In my view the assumed motives for such changes (opacity of markers as supposed by Haspelmath, emphasis or fitting of “an expression in a paradigm into a structural class”, as proposed by Lehmann, 2005:148) are not evident in all of the above cases. The case in § 7.1 (and perhaps the one in §7.3 to a certain extent) more clearly involves a morphologically opaque construction (due to phonological factors) as to certain categories. The rest can be viewed as results of analogical pressure on forms on which the categories seem to be already sufficiently marked. The crucial point is that such processes of “pleonastic affixation” lead to “hypercharacterized” forms which *underline* the categories marked by the new elements *a posteriori*. Joseph (2008:3) in my view points in the right direction when he remarks that “speakers, when innovating, care more about getting appropriate pieces expressed and into the mix, as it were, than they do about observing ordering regularities concerning these elements. This is not to say that anything goes, but recognizes rather that getting the informative pieces into the form is the paramount consideration”. Building on this thought, I would suggest that the above described changes are an instantiation of a tendency to give potentially a separate morphological coding to every morphosyntactic property regardless of the fact that the property is already encoded, albeit cumulatively with other properties by the pre-existent marker. In this process speakers tend perhaps to “spread” the complex morphosyntactic information (tense, person, number, voice) onto more than one element:

-onde [3.person+plural+present+mediopassive] → **-onde-si* [3.person+plural+present+mediopassive] - [3.person+plural]

Limitations on this tendency may be imposed by the length of the resulting construction and the repertoire of available elements. Subsequent phonological and morphological changes (loss of final /i/ due to the trisyllabic window, obsolescence of *-si*) may of course again obscure things:

**-onde-si* > *-undes* (not further segmentable?)

This means that “hypercharacterization” in such cases would be only an epiphenomenon. Furthermore, the constructions referred to in §6 above could fit in this framework as well.

8. Interactions between voices.

8.1. The change of *-ondon* → *-ondan* shows that there are no “watertight” boundaries between voices, and that inflectional forms of one voice can influence the forms of the inflectional paradigm of the other voice, when there exists even a slight semantic relationship between them (cf. above *-ondon* → *-ondan* under the influence of *-an*, though the latter should be regarded as unmarked for voice). In the following case, an inflectional suffix of one voice was adopted as is by the paradigm of the other voice¹⁶. In many varieties of the Aegean Sea, the use of the inflectional suffix of the 1.SG.IMPERF of the mediopassive voice has restored in oxytone active verbs the distinction between the 1.SG and the 3.PL of the imperfect, both of which originally ended in *'-un* (and *'-u* with deletion of final /n/ in some varieties), e.g. **(e)ýélun* ‘I was laughing’, **efórun* ‘I was wearing’ (standard *ýelúsa forúsa*) → *ýelum'na fórum'na* (Kydonies-Lesvos).¹⁷ Interestingly, the mediopassive voice, from which the inflectional suffix originates, is the marked member of the voice system. Furthermore, it was not the unmarked for voice (and hence displaying wider distribution within the verbal system) marker of the 1.SG *-a* that was taken over, in contrast to what happened in many other varieties (e.g. *fórun-α*, standard *forúsa* ‘I was wearing’, see Pantelidis 2008).

8.2. Sporadic attestations of forms like 1.PL. imperfect forms like *stekósame erxósame* (standard *stekómaste/-an, erxómaste/-an* ‘we were standing, we were coming’) (Arcadia, Achaia, see Pantelidis 2006:288), which can be viewed as reanalysis of the forms of the 3.PL. *stek-ósane erx-ósane* as *stekós-ane erxós-ane*, on the basis of the widely used within the verbal system 3.PL.PAST marker *-ane* (active or unmarked for voice). The reanalysis and the spread of the new structure, limited locally to the sub-paradigm of the plural is an interesting evolution, both because it has as a model an ending which is unmarked for voice but which is tacked onto forms which are clearly marked as [+mediopassive], and because of the unexpectedness of the result (no singular forms like **érxos-a, *érxos-es, *érxos-e* are attested), which, as in the previous case, create a new local generalization but an “irregularity” on another level, at least from the specialist’s point of view. The new structure can be subject to alternative interpretations on the part of the linguist (‘new base allomorph *erxós-* + *-ame -ane*’ or ‘base *erx-* + new ending *-ósame*’), all of which could be considered uneconomical and would perhaps go against the perception of the speakers themselves concerning the morphological structure of the inflectional forms. In this case the *-ó-* cannot in my view be considered a marker of tense, as several models of analysis of the Modern Greek verb do. Speakers do not seem to (always) care about the precise status and the precise function of all the elements which make up an inflectional form.¹⁸

¹⁶ In any case, in the mediopassive aorist the inflectional suffixes are identical to those of the active forms of the past.

¹⁷ Kourmoulis 1956:3-4. Further data (from Papadopoulos 1927, Kourmoulis 1956, Katsanis 1995): Samothraki: *bóluman* (orig. *epólun* ‘I was selling’), *xálnuman* (standard *xalúsa*). Imvros: *ayápum, θárrum* (*-um* < **-um'n* < *'-u-mun*). Limnos: *rótum'ne, pirpátum'ne*. Mykonos: *epínun'ne, ezítum'ne*. Kythnos: *a-γápumúne, bórumu*. Krini (Asia Minor): *iyélumúne, irótumúne*. Andros: *ayápumúne, ízjumun* (← **í-zjun* ‘I was living’, standard *zúsa*). Naxos: *zítumun, pérnumun*. Kimolos: *itrávumúne*. Sikinos-Folegandros: *emíljumun*. Rethymno (Crete): *epínunmuné, epérnumuné*.

¹⁸ See also Luschtzky 2000:455. Discussing the issue of what constitutes a morpheme, he points out that while the elements /fl-/ and /gl-/, which appear in many German words with common semantic features, theoretically fulfill the necessary criteria for their recognition as morphemes, such an analysis would be completely unfounded, since the parts of the words that would remain after the segmentation of /fl-/ και /gl-/ (e.g. *-immer-, -irr-, -ucker-, -atter-* etc.) could not be attributed to any morpheme nor could their contribution to the meaning of the whole word be identified. As he himself later on admits, the recognition of a special morphological status for elements such as fl-/ και /gl-/ leads to interesting and justifiable generalizations, which function on

9. Conclusions.

The inflectional paradigm of the mediopassive, and especially of the imperfect, in the Modern Greek dialectal varieties is interesting both for the theory of morphological change and for theoretical notions and issues such as the status of the morpheme, the submorphemic units, the marking of categories and the overall morphological structure. It can, furthermore, be seen within the framework of well-attested cross-linguistic tendencies:

a) Reanalysis shows that speakers often operate on surface forms and are in a constant process of interpretation/analysis of their data.¹⁹ During this process, they are looking for “structure”, especially in longer inflectional forms such as the forms of the mediopassive paradigm (Joseph 1992:131-133, Booij 2007:258). In these forms, they often seem to be looking for overt markers of morphosyntactic categories without necessarily paying special attention to the “correct” linear ordering, to the extent at least that this is deducible from what is known about the morphological structure of the Modern Greek verb.²⁰ In fact, sometimes the function of the elements resulting from the speakers’ alternative analysis of the data is not even clear (Kubrjakova 2000:422,424-425). Thus for example they occasionally analyze linguistic data in a way that goes against the classic synchronic analyses that appear justifiable or transparent from the linguist’s point of view, e.g. the case of *erx-ósane* → reanalysis to *erxós-ane* → 1.PL. *erxós-ame* (beside 1.PL. *erx-ómaste*, alongside singular forms *erx-ómun(a)* *erx-ósun(a)* *erx-ótane* etc). Moreover, they occasionally even go against the basic morphological structure of the Modern Greek verb (as in the case of the reanalysis of the sequences *-mas-* *-sas-* as markers of person+number, and are not always “perfectly” well-founded (at least semantically and syntactically as in this case). As has been remarked, real speakers are far from a “perfect speaker-listener” who has at any moment a grasp of the totality of the system of his language (see Joseph 1992:132-133).

b) Concerning the issue of whether there are constraints on “inter-cell connections” (Joseph 2009:53-54) which might facilitate certain change and render others less probable, there can be no definite and complete answer without a more comprehensive investigation of the changes attested in Modern Greek varieties. A number of tendencies can of course be established. However, it is remarkable that even slots (or cells) which are quite loosely connected with each other (e.g. the same cell in a different voice) may interact.

c) The problems connected with the classic notion of morpheme and the criteria for its identification have already been noted and commented upon in the relevant theoretical literature, and so has the question of its sufficiency for the description and analysis of morphological structure and morphological change. Also, the existence of sub-morphemic units, identifiable on the basis of form and meaning similarities, has been proposed by several scholars (see Luschützky 2000, Kubrjakova 2000). The data from the domain of

the basis of sub-morphemic units. However, it is far from clear whether speakers always attribute (or are even interested in attributing) a clear meaning/function to all the word segments which could be recognizable as units according to specific criteria. See also Bybee 1985:128, Kubrjakova 2000:424-425.

¹⁹ Cf. Booij 2007, p.258: “A [...] source of linguistic innovation besides changing the lexical norm is reanalysis. Language users cannot grasp the system behind a language in a direct fashion. The only evidence they have are outputs of the system, concrete cases of language use. This opens up the possibility that a language user reconstructs the system underlying the perceived outputs in a slightly different way from previous users”. And: “...adult speakers may also change their language through reanalysis, since they are continuously interpreting the outputs that they perceive”.

²⁰ Cf. Joseph 2008, p.3: “...speakers, when innovating, care more about getting appropriate pieces expressed and into the mix, as it were, than they do about observing ordering regularities concerning these elements. This is not to say that anything goes, but recognizes rather that getting the informative pieces into the form is the paramount consideration”.

morphological change in the mediopassive inflectional paradigm (especially the non-perfective past) in Modern Greek varieties seem to confirm, in my opinion, the existence of such elements in inflectional morphology, as speakers seem in several instances to identify within this paradigm units which are bearers of meaning/function but are situated at a sub-morphemic level. Finally, the results of psycholinguistic and neurolinguistic research on speakers' perception of morphological structure can also provide an important and fruitful contribution to the understanding of the mechanisms of morphological change.

References.

- Babiniotis, G. (1972). *To rήμα της Ελληνικής*. University of Athens.
- Booij, G. (2007). *The Grammar of Words, an Introduction to Morphology*. Oxford: Oxford University Press.
- Bybee, J. (1985). *Morphology*. Amsterdam/Philadelphia: John Benjamins.
- Clackson, J. (2007). *Indo-European Linguistics: An Introduction*. Cambridge: Cambridge University Press.
- Coates, R. (2000). "Exponence". In G. Booij et al. (eds.) *Morphologie/Morphology*, Vol. I. Berlin/New York: Walter de Gruyter, 616-630.
- Devlin J. T., Jamison H. L., Matthews P. M., Gonnermann L. M. (2004). "Morphology and the internal structure of words". *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 101.41: 14984-14988.
- Favis, V. (1911). *Γλωσσικά επισκέψεις αναφερόμενοι εις το ιδίωμα Αυλωναρίου και Κονιστρών*. Athens: Paraskevas Leonis.
- Haspelmath, M. (1993). "The diachronic externalization of inflection". *Linguistics* 31: 279-309.
- Horrocks, G. (2010). *Greek: A History of the Language and its Speakers* (2nd edition). Malden et al.: Blackwell.
- Janda, R. D., & Joseph, B. D. (1992). Pseudo-agglutinativity in modern Greek Verb inflection and "Elsewhere". In: *Papers from the 28th Meeting of the Chicago Linguistic Society* 1. Chicago: Chicago Linguistic Society, 251-266.
- Joseph, B.D. (1992). "Diachronic Explanation: Putting Speakers back into the Picture". In G.W.Davis & G.K.Iverson (eds.) *Explanation in Historical Linguistics*. Amsterdam/Philadelphia: John Benjamins, 123-144.
- Joseph, B. D. (2005). "How accommodating of change is grammaticalization? The case of "lateral shifts". *Logos and Language (Journal of General Linguistics and Language Theory)* 6.2: 1-7.
- Joseph, B. D. (2006). "On Connections Between Personal Pronouns and Verbal Endings in the Balkans". In R.A.Rothstein et al. (eds.) *Studia Caroliensia (Papers in Linguistics and Folklore in Honor of Charles E. Gribble)*. Bloomington: Slavica Publishers, 177-188.
- Joseph, B. D. (2008). "What Greek can tell us about the remaking of verb endings". In M.Theodoropoulou (ed.) *Θέρμη και Φως/Licht und Wärme (αφιερωματικός τόμος στη μνήμη του Α-Φ.Χρηστίδη)*. Thessaloniki: Kentro Ellinikis Glossa, 277-281.
- Joseph, B. D. (2009). "Greek Dialectal Evidence for the Role of the Paradigm in inflectional change". *Morphology* 19: 45-57.
- Karanastasis, A. (1982). *Γραμματική των ελληνικών ιδιωμάτων της Κάτω Ιταλίας*. Athens: Academy of Athens.
- Katsanis, G. (1996). *Το γλωσσικό ιδίωμα της Σαμοθράκης*. Thessalonik: Municipality of Samothraki.
- Kollia, A. (1933). "Το γλωσσικόν ιδίωμα της νήσου Κέας". *Αθηνά* 45: 262-285.
- Koefoed G. & J. van Marle. (2004). "Morphological change I: Fundamental issues". In G. Booij et al. (eds.) *Morphologie/Morphology*, Vol.2. Berlin/New York: Walter de Gruyter, 1574-1589.
- Kourmoulis, G. (1956). "Αναλογικά αναμείξεις ενεργητικών και παθητικών καταλήξεων". *Platon* 15: 3-11.
- Kubrjakova, E. (2000). "Submorphemische Einheiten". In G. Booij et al. (eds.) *Morphologie/Morphology*, Vol. I. Berlin/New York: Walter de Gruyter, 417-426.
- Lehmann, C. (2005). "Pleonasm and hypercharacterization". In G. Booij & J. van Marle (eds) *Yearbook of Morphology 2005*. Dordrecht (Netherlands): Springer, 119-154.
- Luschützky, H-C. (2000). "Morphem, Morph und Allomorph". In G. Booij et al. (eds.) *Morphologie/Morphology*, Vol. I. Berlin/New York: Walter de Gruyter, 451-462.
- Mackridge, P. (1990). *Η νεοελληνική γλώσσα, περιγραφική ανάλυση της νεοελληνικής κοινής*. Athens: Patakis.

The Reshaping of the Mediopassive Endings:
Evidence from Modern Greek Varieties

- Minas, K. (1987). "Οι καταλήξεις της οριστικής των μονολεκτικών ρηματικών τύπων της μεσοπαθητικής φωνής". *Dodoni* 16: 21-60.
- Newton, B. (1972). "The Dialect Geography of the Modern Greek Passive Inflections". *Glotta* 50: 262-289.
- Pantelidis, N. (2005). "-όντουσαν". *Meletes gia tin Elliniki Glossa* 25: 473-483.
- Pantelidis, N. (2006). "Reanalysis in Inflectional Morphology: Evidence from Modern Greek Varieties". In: M. Janse et al. (eds.) *Proceedings of the Second International Conference of Modern Greek Dialects and Linguistic Theory (Mytilene, Greece, 30 September-3 October 2004)*. Patras: University of Patras, 282-297.
- Pantelidis, N. (2008). "Η 'ενοποίηση' του παρωχημένου: Ιστορική προσέγγιση με βάση τη μαρτυρία των νεοελληνικών διαλέκτων". In A. Mozer et al. (eds.) *Γλώσσας Χάριν (τόμος αφιερωμένος από τον Τομέα Γλωσσολογίας στον καθηγητή Γεώργιο Μπαμπινιώτη)*. Athens: Ellinika Grammata, 289-302.
- Papadopoulos, A. (1927). *Γραμματική των βορείων ιδιωμάτων της Νέας Ελληνικής Γλώσσας*. Athens: Sakellariou.
- Papadopoulou, M. (2005). "Το ιδίωμα της Πάτμου". *Lexikografikon Deltion* 25: 169-196.
- Ralli, A. (1987). *Éléments de la morphologie du grec moderne: La structure du verbe*. Université de Montréal.
- Ralli, A. (2005). *Μορφολογία*. Athens: Patakis.
- Ruge, H. (1973). "Imperfekt Passiv im Neugriechischen". *Glotta* 51: 142-159.
- Ruge, H. (1984). "Zur Entstehung der neugriechischen Mediopassivendungen". *Folia Neohellenica* 6: 132-143.
- Salvanos, G. (1918). *Μελέτη περί του γλωσσικού ιδιώματος των εν Κερκύρα Αργυράδων*. Athens.
- Taft K. & P. Kougiou. (2004). "The processing of morpheme-like units in monomorphemic words". *Brain and Language* 90: 9-16.