

Brian D. Joseph  
 The Ohio State University  
 bjoseph@ling.ohio-state.edu

## DIALECT EVIDENCE BEARING ON THE DEFINITION OF 'WORD' IN GREEK

### Abstract

Although linguists tacitly operate as if the notion "word" were straightforward, it is in fact one of the trickiest of grammatical constructs to define accurately. Different notions of 'word' may be operative for different levels and/or components of grammar, e.g. a "phonological word" may be different from a "syntactic word" and different from a "lexical word", and different dialects or varieties of a language may differ on the criteria for wordhood and on the status of individual elements. In this paper, accordingly, the issue of how to define "word" for Modern Greek is investigated, with the main emphasis being on how data from various Greek dialects contributes to a pan-Hellenic determination of the tests relevant for identifying which elements are best considered as "words". Crucial to this task for Greek is the analysis of various "little elements", the so-called "clitics" and "particles" that are part of the grammatical apparatus of noun phrases, verb phrases, and sentences, i.e. familiar elements like *tha* and *na*, the weak object pronouns, the negative markers, etc.

### 1. Introduction

Greeks have long shown a fascination with the notion of 'word', as suggested, for instance, by the fact that there were no less than eight terms used for this notion in Ancient Greek, admittedly from different eras and different genres of usage, and thus with different nuances of meaning: *diálektos*, *épos*, *lektón*, *léxis*, *lógos*, *ónoma*, *rhé:ma*, *pho:né:*. Moreover, the very concept of 'word' as a technical construct within Western grammatical theory has its roots in the Hellenistic grammarians' definition, as given in (1) (cf. Robins 1993: 57):

- (1) *léxis estí méros toû katá súntaxin lógou elákhiston*  
 word/NOM is/3SG part/NOM of-the concerning syntax/ACC expression least/NOM  
 'A word is the minimal part of a syntactic construction'

The definition in (1) has withstood the test of time and is the basis, for instance, for the concept in modern linguistic theory of "lexical integrity"; still, it can be updated somewhat to the informal characterization in (2):

- (2) 'word' (informal): unit of organization within a grammatical hierarchy that is above the level of individual sounds and below the level of sentences

and this can be specified further as a more technical characterization, in (3):

- (3) ‘word’ (technical): the output of the morphological component and the smallest unit visible to the syntactic component.

Such a characterization, however, presents at least two problems. First, it is not theory-neutral, and in particular, is not applicable in theoretical frameworks that do not recognize a separate morphological component or which manipulate various grammatical elements in the syntax. Second, different notions of ‘word’ may be operative for different levels and/or components of grammar, e.g. a “phonological word” may be different from a “syntactic word” and different from a “lexical word”. Moreover, with regard to particular languages, one has always to be concerned with internal dialect divisions and how they interact with these notions of ‘word’; it is conceivable that dialects will differ on just how these notions are realized.

Consequently, an approach that many linguists have taken is to work on a language-specific basis, applying various “tests” or “criteria” to look for grammatical generalizations that must make reference to, or allow for the identification of, appropriate higher-level constructs that can be termed ‘word’ in some sense or at some level of analysis, for some dialect or variety.

In this paper, accordingly, I take this approach to the issue of how to identify ‘word’ for Modern Greek, and pay particular attention to how evidence from regional and social dialects bears on this question.

## **2. Grammatical Preliminaries**

As it happens, most of the problematic aspects of deciding about wordhood in Greek focus on various “little elements”, what are often — erroneously or misleadingly — referred to as “clitics”; thus some basic notions about such elements in general are presented in this section, followed in section 3 by a (near-exhaustive) listing of the relevant elements in Greek, with examples, in order to set the stage for a serious consideration of their analysis and the relevant dialect evidence.

The main problem with these elements is that the term most often used to characterize and/or classify them, namely “clitic”, has come to mean for most linguists simply *any* short word-like entity that has some grammatical function and some prosodic deficiencies; moreover, almost no one justifies using the term for any particular element in the language they were looking at, as if it were always self-evident that a given element is a “clitic”

To remedy these shortcomings in the use of the term “clitic”, I take a different view. Following the lead of Arnold Zwicky, I maintain that “clitic” is a most uninformative term; as he puts it (Zwicky 1994: xiii-xv): “clitic ... is an umbrella term, not a genuine category in grammatical theory[; moreover] a variety of phenomena [that] have appeared under the clitic umbrella ... merely have marked properties in one or more components of grammar”.

A similar stance (though taken for somewhat different reasons) concerning the nonutility of the notion “clitic” is to be found in Everett 1996.

In such a view, what must be recognized as morphological/syntactic primitives is not a three-way division of AFFIX vs. CLITIC vs. WORD, but rather simply a bipartite one of AFFIX vs. WORD. What is needed further, though, is the recognition that within each category, there are *typical* (i.e. “core”) and *atypical* (i.e. “marginal” or “marked”) members, but this is required independently of the decision regarding clitics, i.e. even if one were to adopt a basic 3-way affix/clitic/word distinction.

With this simpler inventory of basic elements, the grammar handles (i.e., accounts for or distributes) affixes in the morphology (i.e. the morphological component) and words in the syntax (i.e., the syntactic component). Moreover, elements *must* be designated by the grammar as an affix or a word, i.e. assigned to the morphology or to the syntax; it is one of the functions of the grammar to reflect this status — a putative “cline” that is often posited between these two polar oppositions is merely the linguist’s reflection of the fact that there are typical and atypical members of each category.

As noted above, a useful way of determining where an element falls is by reference to various “tests”, mostly language-specific “behaviors” (though some cross-linguistic universals or tendencies do emerge), that are typical of one or the other type; for the most part, affixes show a greater degree of idiosyncrasy along various parameters (e.g., following Zwicky & Pullum 1983, Zwicky 1985, rigidity in ordering, selectivity in cooccurrence, etc.), whereas words show less idiosyncrasy, inasmuch as they are syntactic entities manipulated by rules of syntactic distribution, which are maximally general (referring to categories only, not to individual lexical items) and which feed directly into semantic interpretation (so that there is compositionality — i.e., a one-to-one mapping — between syntactic rules that build structure and rules of semantic interpretation).

An enumeration of the relevant elements is given in section 3, after which their analysis can be pursued.

### 3. The Range of Relevant “Clitic-like” Elements in Greek

Greek is rather rich in various “little elements” that pose interesting analytic problems. While these have typically been treated as if they were words (in some sense) or “clitics”, some (especially those with grammatical functions) may be analyzable as affixes, possibly inflectional in nature (cf. Joseph 1988, 1989, 1990, 1994, 2000, Forthcoming). A fairly complete listing is given in (4):

- (4) a. elements modifying the verb, clustering obligatorily before it (when they occur), marking:

subjunctive mood: **na** (general irrealis) / **as** (hortative)  
 future (and some modality): **θa**  
 negation: **ðe(n)** (indicative) / **mi(n)** (subjunctive)

- b. elements (generally) correlating with argument structure of verb (“object pronouns”), occurring as closest element to verb (i.e., “inside of” modal etc. modifiers above), positioned before finite verbs and after nonfinite verbs (imperatives and participles); “ACC” stands for direct object markers, “GEN” for indirect object markers:

<u>PERS</u>	<u>SG.ACC</u>	<u>SG.GEN</u>	<u>PL.ACC</u>	<u>PL.GEN</u>
1	<b>me</b>	<b>mu</b>	<b>mas</b>	<b>mas</b>
2	<b>se</b>	<b>su</b>	<b>sas</b>	<b>sas</b>
3M	<b>ton</b>	<b>tu</b>	<b>tus</b>	<b>tus</b>
3F	<b>tin</b>	<b>tis</b>	<b>tis</b>	<b>tus</b>
3N	<b>to</b>	<b>tu</b>	<b>ta</b>	<b>tus</b>

- c. weak 3rd person nominative (subject) markers (with two — and only two — predicates: **ná** ‘(t)here is/are!’ and **pún** ‘where is/are?’), always postpositioned and inseparable from the predicate):

<u>PERS</u>	<u>SG</u>	<u>PL</u>
3M	<b>tos</b>	<b>ti</b>
3F	<b>ti</b>	<b>tes</b>
3N	<b>to</b>	<b>ta</b>

- d. “weakened” (NB ? weak forms, cf. below) nominatives (subject pronouns):

<u>PERS</u>	<u>SG</u>	<u>PL</u>
1	<b>γo</b>	<b>mis</b>
2	<b>si</b>	<b>sis</b>

- e. attitudinal marker (of impatience), **dé**, always phrase-final (except for one fixed expression, **dé ke kalá** ‘with obstinate insistence’)
- f. pronominal marking of possession within noun phrase (so-called “genitive” pronouns, typically occurring at the end of a noun phrase after noun and identical in form with weak indirect object markers but not in all behavioral aspects (see below, §5)):

<u>PERS</u>	<u>SG</u>	<u>PL</u>
1	<b>mu</b>	<b>mas</b>
2	<b>su</b>	<b>sas</b>
3M	<b>tu</b>	<b>tus</b>
3F	<b>tis</b>	<b>tus</b>
3N	<b>tu</b>	<b>tus</b>

- g. definiteness within noun phrase (the so-called “(definite) article”):

<u>CASE</u>	<u>M.SG</u>	<u>F.SG</u>	<u>NTR.SG</u>	<u>M.PL</u>	<u>F.PL</u>	<u>N.PL</u>
NOM	<b>o</b>	<b>i</b>	<b>to</b>	<b>i</b>	<b>i</b>	<b>ta</b>
ACC	<b>ton</b>	<b>tin</b>	<b>to</b>	<b>tus</b>	<b>tis</b>	<b>ta</b>
GEN	<b>tu</b>	<b>tis</b>	<b>tu</b>	<b>ton</b>	<b>ton</b>	<b>ton</b>

- h. the locative/dative preposition **s(e)** ‘to; in; on; at’, always phrase-initial, attaching to whatever occurs next in the noun phrase (but not necessarily always “clitic”; see below)

and examples of each are given in (5), highlighted in **bold**:

- (5) a. **ðe** **ða** **ton** páte s to spíti su  
 NEG FUT him/3SG.ACC go/2PL.PRES to the-house your  
 ‘You won’t take him to your house’
- b. **as** **min** **tus** **ta** púme **ta** néa **mas**  
 SUBJUNC NEG them/GEN them/NTR.ACC say/1PL the-news our  
 ‘Let’s not say our news to them’
- c. **na** **su** éyrafe o jánis  
 SUBJUNC you/GEN write/3SG.PST.IMPFVE the-John/NOM.M.SG  
 ‘John should have written to you’
- d. **pés** **to** **dé**  
 say/IMPV.SG it/ACC de  
 ‘So say it already!’
- e. **pún** **dos?** Ná **tos!**  
 where-is he/WK.NOM here-is he/WK.NOM  
 ‘Where is he? Here he is!’
- f. kséro **γó**  
 know/1SG I/NOM(WKNED)  
 ‘How should I know?’

In what follows, an analysis is given of these elements with respect to how they interact with and shed light on an identification of ‘word’ in Modern Greek, though due to space limitations, attention is focused here primarily on the weak object pronouns (4b) and the genitive possessives (4f); some discussion concerning the other elements can be found in Joseph (1990, 1994, 2000).

#### 4. Toward an Analysis: Different Notions of Word

As noted in Section 1, with regard to (3), it may be that separate notions of ‘word’ need to be recognized for different levels of grammatical analysis. For one thing, there is the notion of ‘grammatical word’, which represents ‘word’ as listed in lexicon (there being nothing in Greek like verb + particle combinations of English), and thus takes in the major syntactic categories (noun, e.g. *spíti* ‘house’; verb, e.g. *lín-* ‘untie’; adjective, e.g. *árostó-* ‘sick’; preposition, e.g. *apó* ‘from’). Some issues that arise relevant to this notion of ‘word’ involve first of all the representation of inflection: perhaps, as with Lyons 1968, what the lexical listing consists of is the **stem** (which might be thought of as thus representing the **lexeme**) and inflected forms (where they exist) are the actual **grammatical words**. Also, the representation of the “little words” of (4) becomes an issue. Many of them have grammatical function (e.g. the elements of the verbal complex) and so they could be **inflection**, properly constituting part of a **grammatical word**. Alternatively, they could be separate **grammatical words** in their own right. Similarly, to the extent that any of these elements, being members of major syntactic categories, are manipulated by the rules of syntax, they can also be considered **syntactic words**.

Another level of analysis in which a separate notion of ‘word’ might be useful concerns the phonology. For instance, the phonotactics of the language could conceivably provide some insight into how to identify a word, if, for instance, there were some combinations of sounds that only occurred word-initially or word-finally (as considered briefly in Joseph 2000). Further, the phonological generalizations in the language might be such as to require reference to an entity that might involve separate grammatical or syntactic words (appropriately defined) that do not behave in ways that are fully phrasal (where a phrase is taken to consist of concatenated independent syntactic words). This depends to some extent on how all the “little elements” with grammatical values are analyzed; if they are inflectional affixes, then much of what might be called a ‘phonological word’ is simply created by regular word-formation and inflectional processes. A promising domain of this sort that other analysts have considered for Greek involves morphophonemics and especially nasal-induced voicing; thus the next section considers this type of evidence for a notion of ‘word’ in Greek.

#### 5. Morphophonemics and the Word, and a Foray into Dialect Evidence

The key morphophonemic alternations in Greek that show interesting interactions with various notions of ‘word’ are those arising from the nasal-induced voicing of the voiceless stops /p t k/. Moreover, a consideration of these alternations requires some recognition of dialect differences in Greek, where “dialect” is to be understood in broad sense, reflecting individual or socially determined varieties, in addition to the more traditional geographic/regional varieties. Furthermore, these alternations lead to an interesting result regarding possessives and weak pronouns, a result that finds support in regional dialect differences.

The basic relevant facts can be summarized as follows (see Arvaniti & Joseph 2000 for

additional discussion and references). Phonologically, the status of the voiced stops [b d g] is tricky. At the lexical level (what corresponds to “grammatical word” in traditional terms), for many (now generally older) speakers, excluding recent loans, [b d g] occur by themselves only word-initially and occur medially only after a nasal; thus **brostá** ‘in front’ but **émboros** ‘merchant’ (not **\*\*éboros**). Moreover, again excluding recent loans, there are no cases word-internally of a nasal + voiceless stop (i.e. no cases of [...VmpV...]). But even for (some) such speakers, the initial stop can sometimes be lightly pre-nasalized in some words, and medially, the preceding nasal consonant can be quite “weak” and sometimes even absent (all subject to a complex of factors including addressee, style, speech rate, etc.), i.e. [mbrostá] / [émboros] ~ [éboros]. Further, for some (mostly younger) speakers, the nasal is (almost) categorically absent. This distribution, even for older speakers, has been somewhat disrupted by loan words, so that again, for *some* speakers, a word such as **robót** ‘robot’ has only a voiced stop (i.e., [...b...] not [...mb...]) and **sampánja** ‘champagne’ shows no medial voicing (i.e., [...mp...] not [...(m)b...]), though for *others*, there can be voicing in such loans and/or borrowed medial voiced stops can be “propped up” with a nasal, giving [sa(m)bánja] and/or [rombot].

There are, however, additional relevant facts: at the phrasal level, involving combinations of some of the “little elements” of (4) with a “host” element, final nasals induce voicing on following voiceless stops at boundaries (and the nasal undergoes place assimilation); e.g.:

- (6)                    /ton patéra/ ‘the father/ACC’ → [tom batéra]  
                           /tin píraksa/ ‘her I-teased’ → [tim bíraksa]  
                           /ðen pírázi/ ‘not it-matters’ → [ðem bírázi]

Optionally (again subject to a complex of factors), in these combinations, the nasal can be weak or even absent, but also, for some speakers, sporadically, there is no voicing whatsoever and sometimes just deletion of the nasal, e.g. [ti(n) píraksa] ‘I-teased her’.

Some linguists have taken the voicing in these combinations as evidence that a level of phonological word must be recognized, combining grammatical/lexical/syntactic words into phrases in which certain phonological effects are located. It is important in this regard to note that the voicing effects, while similar to what is found word-internally, are not identical; for instance, the [ti(n) píraksa] outcome is not found in medial position. Alternatively, if the “little elements” are affixes, one could point to the similarity of the “boundary” phenomena to word-internal combinations with voiced stops, and treat the [ti(n) píraksa] outcome as part of the idiosyncrasy of affixal combinations, thus considering the construct as a **morphological word** or perhaps **morphosyntactic word**, with the affixes as the realization (the “spelling-out”) of various features, such as [+negation] or [+3SG.FEM.DIR.OBJ].

Still, some voicing can be induced by what must be a word in any approach, namely the complementizer **án** ‘if’, as in /án pó/ ‘if I-say’ → [ám bó], for some speakers (maybe only

in fast speech). This fact by itself might tip the balance in favor of the (grammatical-words-combining-into-a-) phonological-word approach and against the affixal/morphological-word approach, except that for some speakers, the usual outcome of /án pó/ is [ám pó], which is definitely not a word-internal type outcome, and in any case it can never become \*\*[a bó], even for speakers who usually do not have a nasal with a voiced stop word-internally. Therefore, there is a real difference between combinations with articles, pronouns, etc. and combinations with more clear-cut grammatical words. While this might be taken by some as evidence for an intermediate construct such as “clitic”, it can just as easily be accommodated in the approach advocated here, as atypical word- or atypical affix-behavior. That is, a word-final -n would not typically trigger voicing on a following stop, but atypically, the word **an** would be such a trigger. Affixes such as the direct object marker **tin** ‘her’ would typically trigger voicing, but idiosyncratically (thus, atypically) could fail to (giving the **ti(n) píraksa** outcome).

There is yet more, however, to consider. The genitive weak pronoun used for marking indirect objects is identical in form to the genitive weak pronoun used for marking possession (cf. (4)), but they show different behavior vis-à-vis nasal-induced voicing. In particular, the object pronouns, which are affix-like in showing idiosyncrasies, high selectivity, strict ordering, etc. (see Joseph 1988, 1989, 1990) are voiced post-verbally after the imperative singular of **káno** ‘do, make’, the only context in which a weak object pronoun occurs after a nasal-final host in the standard language, e.g. /kán tu mja xári/ ‘do for-him a favor’ → [ká(n) du ...]. But the homophonous possessive pronoun **tu** ‘his’ in **ton anθrópon tu** ‘of his men’ (literally, “of-the-men/GEN.PL of-him”) interestingly does not undergo voicing (thus, [...n t...]). Most treatments label both of these as “words” (e.g. “clitic words”), but their differential behavior here is reason for separating them, despite their homophony, and thus under the bipartite division adopted here for treating the object pronouns as affixes and the possessives as words. Relevant here is the fact that the possessives are unaffix-like in being able to move around within the noun phrase; that is, both **o kalós filós tu** ‘the good friend of-him’ and **o kalós tu filós** ‘the good of-him friend’ are acceptable for ‘his good friend’. At the least, however, nasal-induced voicing should probably be separated into a couple of (sub-)processes, and one possible generalization for voicing is that prosodically weak words cannot undergo post-nasal voicing. While one might say that the possessives are thus true clitics, an appeal to typical and atypical behavior for words can work just as well — the prosodic weakness they show would be atypical for a word, but it would give a basis for distinguishing the possessives from the weak indirect object pronouns without recourse to a separate construct of “clitic” or “phonological word”.

Separating possessives from indirect object weak pronouns, as suggested by this nasal-induced voicing evidence, finds support from dialect data. In particular, in the northern dialects of Greek we find ACC for GEN (e.g., 2SG **se** for **su**), for indirect object weak pronouns, e.g. (**se díno** ‘I give to you’, but *not* for possessives, e.g. **ta érya mu** ‘my works’ but *not* \***ta érya me**). Thus, at least at the point of development of the northern dialects, POSS and INDIR.OBJ pronouns, in spite of their identity of form, were kept separate by the



grammar. Their functional differentiation can be invoked here, but it is consistent with their distinct behavior vis-à-vis voicing and is suggestive of their simply being distinct grammatical units, each with its own set of properties despite being homophonous.

### 6. Suprasegmentals and ‘Word’, with More from Dialects

I turn now to another phonological domain, that of suprasegmentals, for the occurrence of stress accent (potentially) bears on definition of ‘word’. Two aspects pertaining to stress are relevant here: the number and placement of the accent.

In general, there is at most a single main stress accent in a grammatical word, underlyingly (in its lexical form), and it must fall on one of the last three syllables. The feminine nouns in *-ia* show all the possibilities: **peripétia** ‘adventure’ vs. **ðimokratía** ‘democracy’ vs. **omorfíá** ‘beauty’. When a clear inflectional suffix is added to a stem, it can trigger a rightward accent shift in a stem that has (lexical) antepenultimate accent, e.g.:

- (7)    ónoma                    ‘name’ (NOM/ACC)  
       onónoma-tos            ‘of a name’ (GEN)

This phenomenon has traditionally been treated as consistent with a principle that the accent in a grammatical word can be no farther from the end of the word than the antepenultimate syllable.

On the other hand, when a pronoun (including the possessives) is added to the **end** of a word with antepenultimate accent, it triggers an accent **addition** on the syllable before the pronoun (and a reduction of antepenultimate accent):

- (8)    to ónoma ‘the name’ / to ðnomá tu ‘the name his’ (i.e., ‘his name’)  
       kítakse! ‘look!’ (IMPV.SG) / kítaksé me ‘look at me!’

This has also traditionally been treated as induced by ban on accent farther from end than antepenultimate syllable (with the reduction triggered by a ban on more than one main stress in a word).

For linguists inclined to treat pronouns as word-like entities of some sort (e.g. “clitics”, with their own maximal projection in the syntax), these facts have motivated a higher level construct such as “prosodic word” (implicit in the accounts of Arvaniti 1991, 1992) or “clitic group” (Nespor & Vogel 1986), or perhaps simply “phonological word”, since the pronouns behave differently from clear affixes (which shift accent, cf. (7)) and from clear word combinations (which have no accentual effect, with each word rather having its own accent). Such a construct could be seen to lend support to the phonological word analysis of nasal-voicing, discussed in §5.

Thus this differential behavior regarding accentual effects on the part of clear affixes, clear

words, and the pronouns *could* provide a basis for distinguishing the weak object and possessive pronouns from “true” affixes.

However, it must be noted that there are several idiosyncratic accentual effects associated with affixes. For instance, the neuter GEN.SG **-tos** provokes placement of accent on second syllable to the left of it; usually, this entails a shift of basic accent position to right by one syllable, as in (7) **ónoma/onómatos**, but with shorter stems, there is no shift, only placement two syllables away’, e.g. ‘verb’ **ríma/rímatos**). Similarly, the neuter GEN.PL marker **-ton** provokes placement of accent on syllable immediately to the left of it; usually this entails a shift of basic accent position to right by two syllables, e.g. ‘name’ **ónoma/onomáton**, but, again, with a shorter stem, this is effected differently, with a rightward shift by one syllable, e.g. ‘verb’ **ríma/rimáton**). Moreover, there are some affixes that are always accented, e.g. the past imperfective marker **-ús-** (as in 1SG **filúsa** ‘I was kissing’), and some that are never accented and provoke no accent shift, e.g. the 1PL **-me** (as in **línome** ‘we are untying’). Therefore, the accent addition with weak pronouns, if they are treated as affixes, could simply be yet one more idiosyncratic accentual effect associated with an affix.

Admittedly, the possessive pronouns also provoke accent addition (cf. (8)), so if one were to say that for this reason they are “clitics” (or atypical, i.e. prosodically special words), one could argue that the weak pronouns should fall into same category. Otherwise, the argument would go, the grammar would have duplication through the multiple statements needed for accent addition, in that some affixes would do it and so would “clitics” (or some words). However, it has already been shown in §5 with regard to nasal-voicing that there are differences (at least for some speakers) between weak pronouns and possessives. Somehow, therefore, these two elements need to be differentiated in the grammar; thus if accent addition with the possessives and weak pronouns is consistent with their both being words, the post-nasal voicing facts are consistent with their each being a different kind of element.

Relevant here is the fact that there are prosodically weak words, in particular the attitudinal marker **dé**, that have different accentual properties. **dé** always “leans” on the end of a host but **never** provokes accent addition: **ðokímase** ‘try!’ (IMPV.SG) / **ðokímase dé** ‘try already!’ (not: \***ðokímase de**). Therefore accentually, **de** and the possessives like **tu** ‘his’ have to be differentiated, so that even within that potential class of elements — let’s call them ‘words’ — accentually distinct behaviors must be stipulated. One could say possessives are “true” clitics, but if accentual behavior is the reason, then presumably the weak pronouns belong in the same class; but what then of the post-nasal voicing differences? Should the grammar recognize **four** (or even more) distinct (basic) morphosyntactic elements: word vs. possessive-type “clitic” vs. weak-pronoun-type “clitic” vs. affix?

My solution here is to recognize only affix and word, and to set some tokens apart within

those categories, by way of recognizing different behaviors and realizing that affixes can show various idiosyncrasies. This may also entail giving up on trying to generalize over accentual behavior as a way of differentiating basic morphosyntactic element **types** (though recognizing differences within larger types). Some words may be atypical accentually, e.g. the possessive pronouns, and others may be accentually normal but prosodically atypical in another way, e.g. the attitudinal marker **dé**. Some affixes are accentually neutral (e.g. 1PL -**me**), presumably the typical case, whereas others provoke various accentual adjustments, e.g. genitive singular -**tos** or genitive plural -**ton** or weak object pronouns.

There is some relevant dialectal data that bears on this analysis. In particular, in various dialects, as discussed by Newton 1972, the same sort of accentual adjustment found with the possessives and the postverbal weak object pronouns can be found with the addition to a stem of some disyllabic forms that ostensibly are affixes. For instance, in Thessalian alongside the 1SG form **érxu-mi** ‘I come’ with, as expected, a single antepenultimate accent, there is the 1PL form **érxu-másti** ‘we come’ with “double” accent similar to the pattern found in (8); such forms contrast with the situation in Standard Modern Greek, where the 1PL is **erxó-maste**, with shifted accent vis-à-vis 1SG **érxo-me**. This double accent in apparently affixal formations is found also in Cretan, and elsewhere. In such dialects, accent “adjustment” in longer forms thus is *not* a basis for distinguishing a class of “clitics” (a situation consistent with affixal analysis advocated here with accent adjustment just one of several possible idiosyncratic accentual effects shown by affixes), unless one takes the potentially circular step of saying that these endings in such dialects have been reanalyzed as “clitics”.

Moreover, there are dialects in which the main word accent is more than three syllables from end. For instance, in Northern Greek dialects in the Crimea, as reported on by Dellopoulos (1977), one finds forms like **timázanandini** ‘they were preparing’ (Urzur-Yalta dialect, cf. Standard Greek **etimázondan**), or in Rhodian, as reported on by Newton (1972) the form **érkumeston** ‘we were coming’ occurs (cf. Standard Greek **erxómastan**). Furthermore, Newton 1972 notes there are dialects with *no* accent adjustment with addition of weak pronouns, e.g. Cypriot **skótose ton** ‘kill him!’ (*not* **skótosé ton**), **to fórema mu** ‘my dress’ (*not* **to fóremá mu**), and states that such forms “occur ... in the standard language”, citing **fére mu to** ‘bring it to me!’ (though **fére mú to** is also possible).

What all this means about accent, in my view, is that it is admittedly a way in which one *might* motivate an affix vs. clitic distinction, or a grammatical word vs. phonological word distinction, but it is not clean, and as long as there is messiness, it is not clear what the benefits are. Moreover, the dialect evidence provides a glimpse of what a form of Greek could be like with regard to accent in longer forms — accent adjustment is not an essential part of being a form of Greek (leaving aside the difficult question of course of what it means to be a “form of Greek”); as such, accent adjustment can be viewed as needing to be stipulated, a position entirely consistent with the analysis adopted here whereby the accentual adjustments are effects found with certain prosodically deficient words (thus with

**tu** ‘his’ but not with the attitudinal marker **dé**) and with some affixes (specifically, the weak object pronominal affixes).

### 7. More on Weak Pronouns — And, More on Dialects

The nasal-induced voicing evidence and the accent evidence show that there is no bar to treating weak pronouns as affixes; there is also some positive evidence based on the criteria for affixhood given by Zwicky & Pullum 1983: *fixed order* (any order different from (5) for these elements is impossible); *selectivity in cooccurrence* (e.g. the weak accusative pronouns are not objects of prepositions, etc.); and, *semantic and morphosyntactic idiosyncrasy*, as shown by the nonreferential and counter-valent occurrence of the weak object pronoun **tin** in an expression such as that in (9):

- (9) **pú** **θa** **tin** **pésu**  
 where FUT her/WK.ACC fall/1PL  
 ‘Where will we go?’ (literally: “\*Where will we fall her?”)

Moreover, optionally, and admittedly with a somewhat marked stylistic status, word-internal placement of the weak pronouns is found in some constructions in Standard Greek, as in (10) (from Athanasios Kakouriotis, personal communication 1988):

- (10) **éxe** - **mú** - **te** **embistosíni** (cf. **éxete** ‘have/IMPV.PL!’)  
 have me/IND.OBJ IMPV.PL faith  
 ‘Have faith in me!’

a phenomenon that points to affixal status for the weak pronouns, since, assuming “lexical integrity”, a word should not be able to be positioned internally with respect to another word, whereas an affix of course could be word-internal. Interestingly, there are dialects more widespread instances of word-internal placement of weak object pronouns. Tzartanos 1909 and Thavoris 1977 have given examples from central Greece, e.g. Thessaly, as in (11), and similar phenomena are reported for Cappadocian in Janse 1998:

- (11) **pémti** ‘(y’all) tell me!’ (as if Standard **pés-mu-te**)  
**δómti** ‘(y’all) give me!’ (as if Standard **δός-mu-te**)  
**fériméti** ‘(y’all) bring me’ (as if Standard **fére-mú-te**)

The **-m-** in the first two forms in (11), originally from the weak ISG pronoun, may now simply be an odd stem extension, since Tzartanos reports that overt objects in more usual positions can co-occur, e.g. **pémti mi ki ména** ‘(y’all) tell *me!*’ (literally “**pémti** me even *me!*”). However, at the point at which the **-m-** first came to occur internally, it presumably had pronominal value and in any case, the fuller form **-me-** in **fériméti** is harder to explain as simply a stem extension. Thus non-word-like placement for the weak pronouns must be accepted as a fact about these pronouns, perhaps most strongly in the regional dialects than in the Standard language, but certainly so for Greek in general.

Finally, there is a further matter pertaining to the positioning of weak pronouns that is relevant to the question of their possible affixal status, and where again, regional dialect facts form an important consideration. In particular, apart from the occasional internal placement seen in (10) and (11), weak pronouns in Standard Greek are generally placed before (i.e., to left of) the verb, but can be after (i.e., to the right of) the verb. Moreover, a few expressions occur in the standard language, e.g. **patís me pató se** ‘hustle-bustle’ (literally “you-step-on me I-step-on you”) with finite verbs but a post-positioned weak pronoun (cf. Mackridge 2000). Some analysts might argue that this differential placement reflects movement, and, since stylistic permutation is characteristic more of words manipulated by the syntax than of affixes manipulated by the morphology, it could be taken as counter-evidence to the affixal analysis advocated here. However, this differential placement is **not** random or stylistically controlled, but rather is for the most part **grammatically** determined: weak pronouns are positioned before **finite** verbs (indicatives and subjunctives) and after **nonfinite** verbs (participles and imperatives), e.g. (cf. Joseph 1978/1990, 1983), as exemplified in (12):

- (12) a. γράψτε το! (\*to γράψτε) ‘Write it!’ (imperative)  
 b. to γράψατε (\*γράφσατε το) ‘You wrote it’ (indicative)

In and of itself, therefore, the facts of (12) could be just a matter of the spelling-out of grammatical feature and not evidence of syntactic movement. And, relatively fixed or frozen expressions like **patís me pató se** could simply reflect some further idiosyncratic behavior in the standard language for the weak pronouns that is consistent with affixal analysis.

It should be noted here that there are some dialects in which weak pronouns are post-verbal with *all* verbs, e.g. Cretan **ákusá ton** ‘I heard him’. It is even the case that some dialects into the modern era, e.g. Cappadocian (Janse 1998), have (or at least had, as did earlier stages of Greek) Wackernagel’s Law positioning of weak object pronouns, with the weak pronouns occurring in second position within appropriate phrasal or clausal domains. Given these differences between Standard Greek and the regional dialects, one might want to argue that they offer a pan-Hellenic argument for the weak pronouns as syntactic entities. Quite to the contrary, though, these differences may simply reflect a different status for these elements in each dialect: Wackernagel-type elements may well be syntactic in nature, still prosodically weak words, for instance. In the case of the Cretan (etc.) vs. Standard language differences, they might indicate different syntactic systems, or maybe, if weak pronouns are a matter of morphology, the differences might be explicable in terms of differences in word-formation processes and especially in spell-out rules for inflectional morphology. It need not be the case, therefore, that the dialect differences tell us anything more than the system that each dialect has, in its own terms, and a priori need not point only to a syntactic treatment of the weak pronouns.

As an aside, it can be noted that these dialect differences provide a basis for rationalizing

the occasional phrases with postposed pronouns on finite verbs in the Standard language; that is, expressions like *patís me pató se* may well be the result of *dialect borrowing*. In that way, the overall set of dialect differences in Greek has contributed to how weak pronouns are to be analyzed in standard Greek by providing some idiosyncrasies of placement in these fixed phrases.

## 8. Conclusion

All of the preceding sections, through their examination of the weak pronouns, one of the key types of “little elements” that pose problems for the identification of the construct “word” in Modern Greek, provide a clear basis for understanding the notion of “wordhood” for this language. Working within a restrictive framework that allows only words and affixes as basic units, and degrees of atypicality within those basic categories, one can account for all the properties shown by combinations of weak pronouns with their verbal hosts, inasmuch as the evidence points towards weak pronouns as being affixes and thus the host-plus-weak-pronoun combinations as being simply words built up in the lexicon via word-formation processes embellished via inflectional processes. To be sure, there is much to be said regarding the other “little elements” that bears on the determination of what a word is in Greek (see, for instance, Joseph 1990 regarding negation) but the foregoing gives an idea of how argumentation concerning those elements could go.

Much of the discussion has also concerned dialects, and this raises two general points. First, some dialects appear to be more “advanced”, so to speak, than the standard language, in that the indications of affixal status for the weak pronouns seem stronger in some dialects than in the standard language and may have been reached earlier chronologically. Thus in a sense, each dialect must be examined on its own terms, even if a pan-Hellenic picture can emerge concerning these elements. This leads therefore to the second point, an interesting meta-question that is more methodological in nature.

In particular, a general question to ask when invoking dialect evidence is whether dialects reveal anything about the standard language. In cases of dialect borrowing, as indicated above, they might provide some relevant evidence for or against a given analysis, but does a comparison of two different dialects tell us how to analyze either of those dialects? Perhaps they can, but only to the extent that an analysis of some arbitrary language, e.g. French, can reveal something about how to analyze some other arbitrary language, e.g. Chinese. Alternatively, one might argue that Dialect A may not be the same as Dialect B, but it is as close to B as *any* speech-form could be, without being A itself, so that the analysis in one can indeed carry over into the other. Moreover, it is clear that some dialects might provide some insight into *where* another dialect might be headed, if, for instance, one dialect possibly shows an extension of what is just below the surface, so to speak, in another. In any case, at the very least, the dialect evidence is interesting in its own right, whatever it might tell us about Greek in general or about the most typically cited type of Greek.

## References

- Arvaniti, Amalia. 1991. *The Phonetics of Modern Greek Rhythm and its Phonological Implications*. Ph.D. Dissertation, University of Cambridge.
- . 1992. "Secondary Stress: Evidence from Modern Greek." In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, ed. G. Doherty and D. R. Ladd, 398-419. Cambridge: Cambridge University Press.
- . & Brian D. Joseph. 2000. "Variation in voiced stop prenasalization in Greek". *Glossologia, A Greek Journal for General and Historical Linguistics* Vol. 11-12, 131-166.
- Dellopoulos, Georgios. 1977. "To vorio idioma sti mesimbrini rosia". *A' Simbosio Glosologias tu vorioeladiku xoru*. Thessaloniki: IMXA, pp. 263-304.
- Everett, Daniel. 1996. *Why there are no Clitics. An Alternative Perspective on Pronominal Allomorphy*. Arlington, TX: Summer Institute of Linguistics.
- Janse, Mark. 1998. "Cappadocian Clitics and the Syntax-Morphology Interface". In B. Joseph, G. Horrocks, I. Philippaki-Warbuton, eds., *Themes in Greek Linguistics II*. Amsterdam: John Benjamins Publishing Co., pp. 257-281
- Joseph, Brian D. 1978/1990. *Morphology and Universals in Syntactic Change: Evidence from Medieval and Modern Greek*. New York: Garland Publishers (Outstanding Dissertations in Linguistics Series) (Expanded and updated version of 1978 Harvard University Ph.D. Dissertation).
- . 1983. *The Synchrony and Diachrony of the Balkan Infinitive: A Study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.
- . 1988. "Pronominal Affixes in Modern Greek: The Case Against Clisis". In D. Brentari et al. (eds.), *Papers from the 24th Regional Meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 203-215.
- . 1989. "I erminia merikon vorion tipon tis prostaktikis kata ti simerini morfologiki theoria". *Eliniki Dialektologia* 1 (1989), pp. 21-26.
- . 1990. "The Benefits of Morphological Classification: On Some Apparently Problematic Clitics in Modern Greek". In W. Dressler, H. Luschützky, O. Pfeiffer, and J. Rennison (eds.), *Contemporary Morphology*. Berlin: Mouton de Gruyter, pp. 171-181.
- . 1994. "On Weak Subjects and Pro-Drop in Greek". In I. Philippaki-Warbuton, K. Nicolaidis, & M. Sifianou (eds.), *Themes in Greek Linguistics (Papers from the First International Conference on Greek Linguistics, Reading, September 1993)*. Amsterdam: John Benjamins Publishing Co., pp. 21-32.
- . 2000. "Towards A Characterization Of 'Word' In Modern Greek (Or, 'Do The Greeks Have An 'It' For 'Word'?)". Paper read at International Workshop On "The Status Of 'Word'", Melbourne, August 2000 [to appear in conference proceedings volume, Cambridge University Press (2002)].
- . Forthcoming. *The Modern Greek Weak Subject Pronoun tos — Its Origins and Implications for Language Change and Language Structure: A Study in Grammatical Change*. To appear in Innsbrucker Beiträge zur Sprachwissenschaft series, Universität

- Innsbruck (2002).
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Mackridge, Peter. 2000. "The Position of the Weak Object Pronoun in Medieval and Modern Greek". Unpublished Oxford University ms.
- Nespor, Marina. & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Newton, Brian. 1972. "The Dialect Geography of Modern Greek Passive Inflections". *Glotta* 50.3-4.262-289.
- Robins, R. H. 1993. *The Byzantine Grammarians. Their place in history (Trends in Linguistics. Studies and Monographs, 70)*. Berlin: Mouton de Gruyter.
- Thavoris, Antonis. 1977. "Morfologika merikon idiomaton tis ditikis Makedonias". *A' Simbosio Glosologias tu vorioeladiku xoru*. Thessaloniki: IMXA, pp. 75-95.
- Tzartanos, Axileas. 1909. "Pragmatia peri tis sinxronu thesalikis dialektu". Athens: Petrakos.
- Zwicky, Arnold M. 1985. "Clitics and particles". *Language* 61.2.283-305.
- , 1994. "What is a clitic?". In Joel A. Nevis, Brian D. Joseph, Dieter Wanner, & Arnold M. Zwicky. *Clitics: A Comprehensive Bibliography 1892-1991. (Amsterdam Studies in the Theory and History of Linguistic Science, Series V: Library & Information Sources in Linguistics, 22)*. Amsterdam: John Benjamins Publishing Co., pp. xii-xx.
- , & Geoffrey K. Pullum. 1983. "Cliticization vs. Inflection: English *n't*". *Language* 59.3.502-513.