7th Mediterranean Morphology Meeting

PROCEEDINGS

7th Mediterranean Morphology Meeting

MORPHOLOGY AND DIACHRONY

EDITORS

Angela Ralli

Geert Booij

Sergio Scalise

Athanasios Karasimos

Nicosia, Cyprus
10-13 September 2009

Patras
2010

Morphology
and **Diachrony**

On-line Proceedings of the
Seventh Mediterranean Morphology Meeting
(MMM7)
Cyprus, 10-13 September 2009

Edited by:

Angela **R**alli, **G**eert **B**ooij, **S**ergio **S**calise and
Athanasios **K**arasimos

E-Proceedings Board

Contents

Foreword

The Mediterranean Morphology Meetings (MMM) are organized by Prof. Geert Booij (Leiden), Prof. Angela Ralli (Patras), and Prof. Sergio Scalise (Bologna). For each meeting they cooperate with a local organizer.

The aim of MMM is to bring together linguists who work on morphology in an informal setting, which guarantees maximal interaction between researchers, and gives young linguists the chance to present their work at a conference of moderate size, where fruitful contacts with senior linguists can be established.

The first seven meetings, in 1997 (Mytilene, Greece), 1999 (Lija, Malta), 2001 (Barcelona, Spain), 2003 (Catania, Sicily), 2005 (Fréjus, France), 2007 (Ithaca, Greece), 2009 (Nicosia, Cyprus) have proven the success of this formula. The attendance was high, many abstracts were submitted, and a number of leading morphologists were invited.

Each MMM has two sessions, one with a specific topic, and another with no topic. The specific topic of the Cyprus meeting was 'Morphology and Diachrony'. Some of the papers of this session are under review for publication in the Journal *Morphology* (Springer). These electronic proceedings (ISSN: 1826-7491) include papers also from the free-topic session.

The permanent scientific committee

Geert Booij
Angela Ralli
Sergio Scalise

Sound laws, inflectional change and the autonomy of morphology

HANS-OLAV ENGER
University of Oslo
h.o.enger@iln.uio.no

1. Introduction

This paper addresses a familiar issue; reduction of case in a European language. Old Norse, which is an idealized version of the language that we suppose to have been spoken on Iceland, the Faroe Isles and in Norway around 1200, had four cases on nouns, as Table 1 illustrates:

Table 1: Two Old Norse masculines, indefinite singular only

	ARMR	GRANNI
Nom	armr	granni
Acc	arm	granna
Dat	armi	granna
Gen	arms	granna

By contrast, most Scandinavian dialects today have no case opposition on nouns. So far, data have been simplified, but not distorted; Mørck (2005: 1130) says that "T*he central theme in the history of the nouns is the loss of case inflexion in the Mainl. Scand. languages*". So the question is logical: Why – and how – does the change from four cases to none happen? If there is anybody who thinks that we already know the answer, and that phonology is all, they are in for a surprise.[1] My main claim is that morphology must have some independent role to play.

The rest of this paper is structured as follows. Section 2 gives an outline of a relatively widespread account of the morphological change, in which phonology is "the prime mover". In section 3, we look at some arguments against this traditional and essentially Neogrammarian account. Section 4 is shorter, and the topic is not case, but gender. In some dialects, the number of genders has been reduced from three to two. This change is perhaps not as fully understood as the loss of case, but I suggest that a purely phonological account is not entirely satisfactory for gender, either. In the final section, some theoretical implications are suggested.

2. The traditional explanation

The question why case is lost has been addressed many times before. A widespread view is that the loss of the case opposition is triggered by phonology. Compare the following

---

[1] Thanks to the audience at the MMM 7 in Nicosia, and to Arne Torp and Andrew Carstairs-McCarthy.

quotation from a contribution to a handbook by Delsing (2002: 939), who addresses the closely related languages – or dialects – Old Danish and Old Swedish:[2]

> "*During this period [1100–1350] the morphological system of ODan [Old Danish] and OSw [Old Swedish] started on its way towards the simpler [NB] modern system.* **The causes of this simplification are to be found primarily within the phonological system.** *The weakening of unstressed vowels in suffixes and the dropping of final –R [corresponding to the suffix -r in Table 1 above, HOE], which starts during this period, beginning in ODan., reduce many case, gender and number distinctions in the nominal system.* **These changes are accompanied by analogical processes internal to the morphological system** *[...] The true loss of case and gender distinctions belongs to a later period, but the changes during this period are crucial to the ensuing changes*" p. 939 [emphasis added here, HOE]. Delsing also says that "*weakening of vowels in unstressed syllables causes severe reductions in the morphological system*" ... "*All gender and case distinctions disappear among the weak nouns in the sg. [...] In the weak adjectival paradigm all case, number and gender distinctions are lost*" p. 937.

The reasoning is not always made quite as explicit as here, but that is not particularly surprising, within the vexed field of 'explanations' in diachronic linguistics. Let us first look at two examples where phonology would seem to provide us with sufficient explanation; I just repeat Delsing's points. The opposition between nominative *armr* and accusative *arm* can be lost due to what one may call *r*-deletion. There is evidence for this *r*-deletion in other categories in Danish, notably the plural of nouns.[3] Similarly, the opposition between nominative *granni* and oblique *granna* can be lost due to "neutralisation" or "vowel weakening". All word-final (unstressed) Old Danish vowels /i, a, u/ come out as *e* in Modern Danish. Considering that the formal opposition between *granni* and *granna* shows up in the final vowel, it is not terribly surprising that this "weakening" should have consequences.

The idea, in both cases, is that the phonological change removes so many of the relevant exponents that the case opposition comes to be something of a lexical irregularity, something that is so numerically weak that it cannot be upheld, and is lost by analogical means. I do not think anybody has ever suggested any numerical limit as to how many per cent of the nouns, which have to express a certain opposition for the opposition to prevail. But there is already considerable syncretism elsewhere in the nouns. Old Norse feminines and neuters given in Table 2 illustrate this:

Table 2: Two Old Norse feminines, indefinite singular only

---

[2] The relation between Old Norse, Old Danish and Old Swedish is tricky, but we cannot go into the issue here.

[3] Though only after vowels. Thus, what in Old Norse would be *bǿkr* 'books' is, in Modern Danish, *bøger*, with no deletion. This indicates that the deletion account is more problematic than it seems at first. We return to this point below.

	BYGÐ	LAND
Nom	bygð	land
Acc	bygð	land
Dat	bygð	landi
Gen	bygðar	lands

There are also some masculines (e.g. NAGL) that do not evince any formal differentiation between the nominative and the accusative. And some do not evince any formal differentiation between the dative and the accusative (GESTR). Some few types of non-masculine nouns do not evince any case distinctions in the singular at all (fem. ELLI, neut. HJARTA). (See e.g. Haugen 1995 for further detail.)

Now, the idea that phonology must ultimately be responsible for the case loss seems appealing, for at least three reasons:

1) The morphological change and the phonological change are close in time. By the Neogrammarian account, they are not two independent innovations that just happen to take place at almost the same time. They are related in a way that makes sense. Phonological reduction is also a process that is not only relevant for suffixes; it happens with a host of consonants, and appears to be more general.

2) The case opposition is expressed mainly by suffixes in Old Norse, as are indeed most inflectional oppositions in that language, and the tendency for phonological change to eliminate word-final syllables is familiar in North Germanic, which has become less "inflectional". The changes from Old Norse to modern Norwegian hence fit in a larger picture; unstressed syllables and inflectional suffixes have been dropped before.

3) This is a fairly traditional Neogrammarian scenario (and hence well-established) – phonology is seen as the 'prime mover', morphology as merely 'reactive'. This may fit with our picture of *Lautgesetze* vs. analogy in diachronic linguistics in general.

The theoretical view outlined above is perhaps particularly linked to the Neogrammarians, but it is, I would emphasise, accepted – at least in part – up to the present day.[4] The quote from Delsing illustrates this for Old Swedish and Old Danish, and similar views are expressed for other languages. Compare Trask (1996: 128):

"*The elaborate case-systems of Latin and Old English depended crucially upon distinctions in the final syllables of inflected nouns; as phonological changes began to reduce and to obliterate those final syllables, prepositions came to be used more frequently to reinforce the case distinctions which were beginning to get lost; the increased use of prepositions made the case-endings less significant than previously, and so there was less reason to hang onto the remaining case-endings. Consequently, English and the modern descendants of Latin, like Spanish, French, and Italian, have lost their earlier case-systems completely (except in a few pronouns), and replaced them with analytical (isolating) constructions involving prepositions*".

The very same words can be read in the 2007 edition.

---

[4] Also Wetås (2008: 22) notes that in the literature, the loss of case inflexion is often presented as primarily phonological changes with secondary morphological consequences.

3. Against the traditional Neogrammarian scenario for case

A number of problems with the traditional Neogrammarian conception outlined in 2 have been pointed out in the Norwegian literature. I shall survey these problems, drawing heavily on research carried out by previous scholars.

3.1. The genitive suffix –*s*

To begin with, nobody has ever claimed that the main genitive suffix -*s* is lost due to sound law. There simply is no diachronic process of *s*-deletion, equivalent to that of *r*-deletion, in Norwegian, nor in other North Germanic languages[5]. Nevertheless, the genitive case is lost even in Faroese, which is a fairly conservative variety, retaining the three other cases. It is only in Icelandic that the genitive suffix -*s* has remained fairly unchanged. Many Norwegian dialects have got rid of the genitive -*s* entirely. Admittedly, there are a few Norwegian dialects (and many Swedish ones) that retain an element -*s* indicating possession up to the present day (see e.g. Torp 1973, Gunleifsen 2009), but this element does not behave grammatically as Old Norse -*s* did. Its present status is a matter of much dispute (see e.g. Börjars 2003, Norde 2006 and further references therein). My point is only that there are too many dialects in which it seems implausible to operate with a general phonological rule of *s*-deletion – and still they have no genitive -*s* today.

　　This is all the more noteworthy, because in late Old Norse, after 1200, the genitive marker -*s* was particularly productive (as noted by both Wetås 2008 and Knudsen 1967); just as it was in late Old Swedish (Wessén 1969: 136). Thus, it spread to masculines that, originally, did not have it (cf. late Old Norse *sons* vs. the older *sonar*) and even to feminines. In other words, in the late mediaeval language, the suffix -*s* is what Natural Morphologists refer to as a "super-stable marker". Such markers are characterised by spreading even though no associated class change occurs (Wurzel 1984: 139). That is, even if masculines from other declensions than ARMR get the genitive suffix -*s*, which they previously did not have, they do not join the paradigm wholesale. Interestingly, Dammel & Nübling (2006) suggest that a super-stable marker is indicative of a breakdown in the inflexional system. They use Scandinavian -*s* as one of several examples.

3.2. The definite dative

So far, we have focussed on the indefinite singular. There is, however, also a definite singular. Compare Table 3.

Table 3: The definite singular of two masculines

Nom	armrinn	granninn
Acc	arminn	grannann

---

[5] Already Knudsen (1967) argues that the loss of -*s* is not plausibly attributed only to phonology.

Dat	arminum	grannanum
Gen	armsins	grannans

By traditional accounts (e.g. Enger 1993, Haugen 1995), the opposition between indefinite and definite is inflectional in Old Norse; the exponent of definite is thus an inflectional suffix.[6] So far, we have seen that the case inflection as a rule is lost in the indefinite. In the definite, however, the dative has stood its ground better – in fact, up to the present day in many dialects, while the nominative was by and large lost by 1500. To be sure, the definite dative is being lost today (see e.g. Sandøy 2000), but that is another story; the point is that it stayed on for so long – and its geographical distribution. Venås (1993: 262) notes as an "interesting geographical aspect" of the case reduction that "the dative is lost almost everywhere in that area that perhaps has retained the Old Norse phonological structure best, South-West Norway" [my translation]. So, in the South-West, /e/ and /a/ do not merge, most old /a/s are faithfully retained, quite unlike Danish. If what triggers the morphological change – dative loss – is the change away from Old Norse phonological structure, one would expect there to be less such loss in the area where the old phonological structure is best retained; but this is simply not what we find.

Making the case for phonology even worse, Knudsen (1967: 12ff) points out that the dative case has been retained in many Norwegian dialects that in general have had much phonological reduction in the final syllable. So the definite dative has been retained better where one might expect it to be lost on purely phonological grounds.

Knudsen (1967:13) argues that the retention of the dative "*is not primarily decided by the effects of the sound laws, but follows other lines: a tendency to retain the dative in inland dialects that preserve the older system with richer formal categories [his example is inland East Norway]; loss of dative in coastal (and town) dialects where more traffic […] has given the preconditions for a language with less formal distinctions*" [my translation].

This brings up the sociolinguistic factors. Contact apparently plays a role, as noted by several Scandinavian scholars (Knudsen 1967, Wessén 1967: 185, Torp & Vikør 2003). I have already quoted Trask's *Historical Linguistics* above. In the more recent 2007 edition, these lines remain the same (p. 159), but Robert McColl Millar has added an extra sentence: "*Many would claim, however, that language contact at least encouraged the developments involved*". These are words of wisdom, and another example in favour of this view will be brought up in the discussion of gender in section 4 below. But even when the factor of contact or sociolinguistics has been added, the phonological account does not quite hold, as the demise of the Faroese genitive makes clear (cf. 3.1 above). It is hardly plausible to attribute this development to contact.

3.3. How phonological is it really?

---

[6] According to Faarlund (2009), it should be considered cliticisation, but this issue is a large one and not essential for present purposes.

In the very same volume as Delsing, Wurzel (2002: 258) questions – or rather dismisses – the phonological motivation for the loss of case:

> "*The first stage of this development [loss of formal distinction nominative-accusative in Swedish] sees the strong masculine forms of the type OSw. nom.sg. hund-er – acc.sg. hund [dog] losing the nominative marker -er.* **This cannot be a phonological reduction**, *as the change affects neither stem-internal -er, as in OSw. biter, viter > Mod.Sw. bitter, vitter ['bitter', 'wise', both adjectives, HOE], nor the plural marker -er as in OSw. bok – nom.pl. bök-er, nat – nom.pl. nät-er > Mod.Sw. bok – böcker, natt – nätt-er [book, night]*".

Recall from section 2 above that Delsing used Old Danish to make his point, and *r*-deletion is a better candidate for a regular phonological change in Danish than in Swedish. In fact, however, *r*-deletion probably is not a regular phonological change in Danish, either, for two reasons. Firstly, in the verbs, *-r* is not deleted. If an *-r* is deleted as a plural marker, while the homophonous present tense marker is not, this is not a regular phonological change. Secondly, in nouns, the plural marker *-r* is apparently deleted depending upon declension (sic!). Thus, while Danish has *hæste* 'horses' with deletion, it has *gæster* 'guests' without; the reason being that the latter noun belonged to the Old Danish *ir*-declension, the former to the *ær*-declension.

But let us return to Wurzel's point: In Swedish, the deletion process cannot be purely phonological. Norwegian dialects are also interesting here: There are some (e.g. Western) dialects that also have *r*-deletion in the plural of nouns and elsewhere in verbs, and this would at first seem to fit the phonological picture, just as in Danish. But in some other dialects (mainly South-Eastern), *r*-deletion is neither attested in the plural nor in the verbs, so these dialects are essentially like Swedish, as described by Wurzel here. The problem for the phonological account is that the nominative suffix is not retained any better in dialects that do not display (phonological) *r*-deletion. If *r*-deletion is not squarely phonological in all those dialects where case is lost, then *r*-deletion cannot be a necessary condition for the loss of case.

So both with the nominative suffix *-r* and the genitive suffix *-s*, a phonological account seems to lead to at least partly wrong expectations. The same seems to be the case for the dative suffix rendered as *-i* in Old Norse in Table 1. Barðdal (2009) observes that its cognate, the dative /e/, and a verbal suffix /e/ behave very differently in Swedish. Barðdal concludes that the motivating factor for the case loss is squarely non-phonological.

In a fairly recent thesis, Wetås (2008) also presents important arguments against the idea that the loss of case is only phonologically triggered. On the basis of an empirical study of charters from West Telemark, she argues that;

> 1) the change in case inflection behaves differently for proper nouns than for common nouns; proper nouns lose case inflection earlier
>
> 2) morphologically complex proper names behave differently from simplex proper names; case is lost earlier with complexes

> 3) the case reduction behaves differently for person names than for place-names; case is lost earlier with person names

Wetås suggests that all three observations follow from a parameter we may call 'naminess' (or propriality): In general, person names are more 'name-y' than are place-names, morphologically complex proper names are more name-y than are simplex names.

So far, I have quoted others. Finally, let me add an observation of my own, which has to do with epenthetic vowels. In many varieties of Scandinavian, an epenthetic vowel *-e* is inserted quite early before *r*. Thus, we find, in certain West Nw. dialects, for example, a change from *armr* to *armer*. It is not quite clear (at least not to me) exactly when this epenthesis takes place, but the 13th century seems to be a reasonable guess. The loss of the epenthetic vowel cannot, at least not in all the dialects, be attributed to regular phonological change: In the dialect of Jæren, in the South West, the epenthetic vowel is retained in the adjectives (as in the verbs), but not in the nouns. We find, for example, *ein sjuge hund* 'a sick dog' < late Old Norse *ein sjuge hunder.* The epenthetic vowel is presumably not lost by regular phonological change, since it respects the grammatical categories. In other words, there is no phonological reason why this dialect does not have *hunde*. But then, *hunde* is just as clearly a nominative as *hundr*. Jæren is probably representative of most West Norwegian on this point.

There are also Norwegian dialects that have retained the epenthetic vowel even in the nouns (West Telemark, Setesdal, Sunnfjord, parts of Nordfjord, parts of Sunnmøre, cf. Skjekkeland 2005: 62), so that we find *hunde, arme*, going back to the old nominative. The formal opposition to the accusative could thus easily stay on, phonologically speaking. But none of these dialects have retained the case system.

3.4. Phonology cannot be all: Knudsen (1967)

The upshot so far is that phonology can hardly be all. This is also the conclusion of Knudsen (1967), who argues explicitly against the Neogrammarian scenario. He says it may be too simple to assume that the sound laws alone decide: The "sound laws alone are as a rule not capable of making old grammatical categories disappear, although they certainly can speed up the development to a very high degree" (my translation).

4. Gender

We now leave the issue of case, and turn to another issue, that of gender. This is a different, but related topic, in that in many dialects of Scandinavian, the masculine and the feminine "merge". Again, one may wonder why.[7] And again, I shall try to argue that phonology does not quite suffice.

For illustration, let us consider a recent paper by Duke (2009). This paper is in fact so good that it is unfair to use it here, but even Duke, to my mind, illustrates what might be

---

[7] This territory is less well charted than that for case. But a familiar example is from the Bergen dialect, where Jahr (e.g. 2009) and Nesse (2002) take the merger to be due to contact; Perridon (2003) does not.

called "the Neogrammarian bias". Her focus is on gender in Swedish, which – in the standard language – has been reduced from a three-way opposition to a two-way opposition. For the adjectives, Duke suggests that if the final syllable, i.e. the epenthetic vowel and /r/ be lost from the Old Swedish nominative masculine singular *langer*, then the difference between the masculine and the feminine is thereby considerably weakened. The implication is that phonology accounts for the merger.

To be sure, Duke advances her argument for Old Swedish. If we try to transfer it to the next dialect down the road, i.e., Old Norse, the argument turns out to be problematic, however. The important empirical argument against transferring Duke's idea to Old Norse can be found in the dialects. In the Jæren dialect of Norwegian, the final /r/ has been lost from adjectives, and there is indeed a phonological rule of general *r*-deletion in this dialect. The epenthetic vowel is retained in the adjectives, but not in the nouns (cf. 3.3 above). The opposition between the feminine and the masculine is retained in many other categories, but the originally masculine adjective *sjuge, lange* is transferred analogically to the feminine adjectives as well. Thus, when the formal differentiation between feminine and masculine adjectives is lost, this has absolutely nothing to do with loss of phonological marking, only with analogy from the masculine. In this dialect, then, we find, just as in Modern Swedish, there is syncretism between masculine and feminine of the adjectives.[8] But the point is that the masculine-feminine merger comes about by other means than phonology. Furthermore, the loss of gender opposition in the plural, which is general in Scandinavian, simply cannot be accounted for by phonology alone (Enger 2010).

Note also that in the Romerike dialect of East Norwegian, unlike Jæren, there is no general phonological *r*-deletion. This is shown by the present of strong verbs (*kommer*), but also by some adjectives, such as, traditionally, *blår, bakketer* (cf. Refsum 1954). Again, the traditional masculine form is transferred to the feminines – for these adjectives. Thus, there is a masculine – feminine merger here too, and phonology cannot be the reason.

5. Conclusions

I suggest several conclusions:

1) Phonology cannot be the sole prime mover

I am not saying that phonology has no role to play.[9] When the dative is retained better in the definite singular than in the indefinite (cf. 3.2), this must have to do with the phonological difference – the simple observation that there was "more" phonological expression in the definite (e.g. Venås 1993: 262). The indefinite is otherwise usually taken to be unmarked in comparison to the definite, so it is hard to see any purely morphological account for that difference. But it seems impossible to believe that phonology is the only prime mover. There is nothing revolutionary in this conclusion,

---

[8] Though this does not happen for so many adjectives in the Jæren dialect as in Swedish.
[9] On this point, my account differs from that of Barðdal (2009).

which may be rather trite to some specialists in Norwegian, cf. the following quotations from Torp & Vikør (2003: 96):

> "...*But this much is certain:* **Phonological development cannot alone be responsible** *for the fact that most of the Old Norse morphology has been lost on the way towards Modern Norwegian, for there are very many forms that ought to have been different in Modern Norwegian if sound laws alone had had control.* **Perhaps the most plausible hypothesis is that several tendencies have worked simultaneously**: *Phonological development has no doubt worked so that a good many different forms had to merge. But at the same time, we find innumerable examples that the same sounds develop entirely differently in different grammatical surroundings.... The full and complete answer as to why exactly these changes happened [...] may never be found. But partial insight is not to be scorned*" Torp & Vikør, p. 96f. [emphasis added here].

In a similar vein, also Venås (1993/1971: 262) holds both morphology and phonology responsible for the case loss. In this picture, morphology is not merely reactive.

The Neogrammarian model has been tremendously influential in diachrony (cf. Bynon 1977: 15), and for good reasons, to be sure. Yet the over-reliance on phonological accounts may be a mistake.

2) Multiple motivation

A reasonable account of the loss of case seems to involve 'multiple motivation'. Phonology cannot be all, neither can contact (compare e.g. the Faroese genitive). We should beware of the 'exclusionary fallacy', as Langacker (1987) reminds us: "*The gist of this fallacy is that one analysis, motivation, categorization, cause, function or explanation necessarily precludes another. From a broad, pre-theoretical perspective, this assumption is gratuitous and in fact rather dubious, in view of what we know about the multiplicity of interacting synchronic and diachronic factors*" (p. 28) "*Finally, it can be noted that diachronic questions are often posed in dichotomous terms [...] Did construction C arise internally or was it borrowed from a neighboring language? By now most scholars recognize that questions like these are simplistic*" (p. 30). This is explicitly recognised in the quotation from Torp & Vikør above.

The change from four cases to none is one of perhaps three classical chestnuts studied over and over again in Norwegian diachrony, by generations of scholars. If we are left we a partial account of multiple motivations here, this suggests (at least to me) that multiple motivation is plausible in general.

3) The value of dialectal evidence

A methodological point that I hope to have shown is the value in consulting dialectal evidence, simply because our data base is broadened (cf. also Harnisch 2000, Nübling 2005, Ralli 2009). As long as we consult one standard language only, say Danish, the putative relation between phonological and morphological change seems reasonable enough. But it is well known that standard languages are also '*the least interesting kind of language for anyone interested in the nature of human language*' (Hudson 1996: 34). It is

also well known that the distinction between language and dialect is quite problematic in a Scandinavian context; thus, Norwegian as spoken in Oslo is probably more intelligible to a Stockholm speaker of Swedish than is a conservative Northern Swedish dialect. It has been necessary to look into details in a number of dialects in order to show what is wrong with the postulated diachronic connection between phonological loss and case loss.

4) Our view of morphology

By a standard Neogrammarian view, morphology does not really change "by itself", it changes in response to changes elsewhere. Morphology is merely "reactive". If, however, morphology really – at least to some extent – is "by itself" (Aronoff 1994, Carstairs-McCarthy 1994, 2001, Maiden 2004, 2005), then one would not expect the classical and essentially Neogrammarian scenario to be *the whole* truth. And in fact, it is not.

Other authors have also advanced accounts of a more purely morphological kind (Barðdal 2009, Wurzel 2002). It remains to be discussed in further detail how convincing these accounts really are. My main point today has been to go against the more familiar, Neogrammarian scenario.

In synchronic studies, morphological phenomena have over and over been described as 'really' phonology; but in recent years, we have come to recognise the aprioristic assumption that morphology should be reduced to phonology as mistaken (e.g. Lass 1984, Comrie 1986, Carstairs 1988, Anderson 1992, 2008). We should beware of subjugating morphology to phonology (Maiden 2009). Perhaps we need to think more about this in relation to diachrony as well; also in a diachronic perspective, there is some autonomy to the morphological component.

5) What about other languages?

The conclusion we have reached on Norwegian diachrony opens for the possibility that the phonological account is not really entirely adequate for other Germanic or Romance languages, contrary to what Trask implies. For German dialects, Nübling (2008: 313) observes that although the case loss is partly due to phonological change, the phonological developments do not imply that the morphology is unable to resist phonology (see also her p. 322). In contemporary Standard German, Harnisch & Nübling (2004: 1906) observe a tendency to use proper names uninflected, that is, in the nominate even where the syntactic context would seem to require something else (*in einem Bericht des "Neuer Tag"* instead of *des "Neuen Tags"*). This corresponds to the relevance of 'naminess' that Wetås found for Norwegian (cf. 3.3 above). This factor appears to be relevant also in French; Schøsler (2001) reports a clear tendency for nouns that denote humans to retain their old declension longer than nouns denoting non-humans – "[p]roper nouns, however, are an exception: they lost their case marking early" (p. 172). Also Barðdal (2009) questions the phonological account for Germanic.

Obviously, case loss in Germanic and Romance is a large question, and it cannot be answered here. Nevertheless, observations like those made by Nübling, Harnisch & Nübling, Schøsler and Barðdal may make us wonder if the phonological account really is adequate.

Envoi

One of the major virtues of the Neogrammarian account is that the morphological change does not emerge as an independent, isolated observation; rather, it follows from something else, with which it is connected. So, an obvious drawback with the account advocated here is that the morphological change no longer necessarily had to follow from something else, and that it is one out of several changes. In one sense, the phonological account is so much more elegant, so much more appealing than what I have suggested here. The only defense I can offer here is to quote Anderson (1992: 346):

"it is important not to let one's æsthetics interfere with the appreciation of fact."

The Scandinavian case and gender reduction can hardly be attributed only to phonology.

References

Anderson, S.R. 1992, *A-morphous morphology,* Cambridge University Press, Cambridge.

Aronoff, M. 1994, *Morphology by itself,* MIT Press, Cambridge, Mass.

Barðdal, J. 2009, The development of case in Germanic. Barðdal, J. & Chelliah, S.L. (eds), *The Role of Semantic, Pragmatic and Discourse Factors in the Development of Case,* John Benjamins, Amsterdam.

Börjars, K. 2003, Morphological status and (de)grammaticalisation: the Swedish possessive. *Nordic Journal of Linguistics* 26, Cambridge University Press, Cambridge, pp. 133-163.

Carstairs, A. 1998, Some implications of phonologically conditioned suppletion. *Yearbook of Morphology 1988*, Foris, Dordrecht, 1988, pp. 67-95.

Carstairs-McCarthy, A. 1994, Inflection classes, gender and the Principle of Contrast. *Language* 70, Linguistic Society of America, Baltimore, pp.737-788.

Carstairs-McCarthy, A. 2001, Umlaut as signans and signatum, *Yearbook of Morphology 1999*, Kluwer, Dordrecht, pp. 1-23.

Comrie, B. 1986, The Maltese Pharyngeal. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39, Akademie-Verlag, Berlin, pp. 12-18.

Dammel, A., Nübling, D. 2006, The Superstable Marker as an Indicator of Categorial Weakness? *Folia Linguistica* XL, Mouton de Gruyter, Berlin, pp. 97-113.

Delsing, L.-O. 2002, The morphology of Old Nordic II: Old Sw. and Old Dan. Bandle, O. et al. (eds), *The Nordic Languages*, *volume I.* Mouton de Gruyter, Berlin, pp. 925-40.

Duke, J. 2010, Gender reduction and loss in Germanic: three case studies. Nübling, D., Dammel, A., Kürschner, S. (eds.), *Kontrastive Germanistische Linguistik*. Olms, Hildesheim

Enger, H.O. 1993, Morfemet i beskrivelsen av bøyingsmorfologien til norrøne substantiv – utfordring eller hindring? *Maal og Minne*, Det norske samlaget, Oslo, pp. 12-28.

Enger, H.O. 2010, Gender and contact: a Natural Morphology perspective. Siemund, P. (ed), *Universals and Variation.* Mouton de Gruyter, Berlin.

Faarlund, J.T. 2009, On the history of definiteness marking in Scandinavian, *Journal of Linguistics* 45, Cambridge University Press, Cambridge, pp. 617-639.

Gunleifsen, E. 2009, *Prototypiske attributive possessive uttrykk*, Ph.D. thesis, University of Agder, Kristiansand, 2009.

Harnisch, R. 2000, Morphologische Theorie und dialektale Empirie. *Sprachwissenschaft* 25, Carl Winter Universitätsverlag, pp. 367-386.

Harnisch, R., Nübling, D. 2004, Namenkunde. Booij, G.E. et al. (eds): *Morphologie/Morphology,* vol. 2, Mouton de Gruyter, Berlin, pp. 1901-1910.

Haugen, O.E. 1995, *Grunnbok i norrønt språk*, Gyldendal Ad Notam, Oslo.

Hudson, R.A. 1996, *Sociolinguistics*, 2nd edn., Cambridge University Press, Cambridge.

Jahr, E.H. 2009, *Språkhistorie og språkkontakt*, Novus, Oslo.

Knudsen, T. 1967, *Kasuslære*, Universitetsforlaget, Oslo.

Kusters, W. 2000, Morphological simplification: more than erosion? Gilbers, D., et al. (eds), *Languages in Contact*. Rodopi, Amsterdam, pp. 225-231.

Langacker, R.W. 1987, *Foundations of Cognitive Grammar, volume 1*, Stanford University Press, Stanford.

Lass, R. 1984, *Phonology*, Cambridge University Press, Cambridge.

Maiden, M. 2004, Into the past: Morphological change in the dying years of Dalmatian. *Diachronica* 21, John Benjamins, Amsterdam, pp. 85-111.

Maiden, M. 2005, Morphological autonomy and diachrony. *Yearbook of Morphology 2004*, Springer, Dordrecht, pp. 137-175.

Maiden, M. 2009, 'Suppletion in diachrony: a new case-study', paper read at the MMM conference Nicosia.

Mørck, E. 2005, Morphological developments from Old Nordic to Early Modern Nordic. Bandle, O. et al. (eds), *The Nordic Languages, volume 2,* pp. 1128-1149.

Nesse, A. 2002, *Språkkontakt mellom norsk og tysk i hansatidens Bergen*, Novus, Oslo.

Norde, M. 2006. Demarcating degrammaticalization: the Swedish s-genitive revisited. *Nordic Journal of Linguistics* 29, Cambridge University Press, Cambridge, pp. 201-239.

Nübling, D. 2005, Forschungsperspektive zur Nominalmorphologie deutscher Dialekte. Eggers, E., et al. (eds), *Moderne Dialekte – Neue Dialektologie*, Franz Steiner Verlag, Stuttgart, pp. 45–87.

Nübling, D. 2008, Was tun mit Flexionsklassen? Deklinationsklassen und Ihr Wandel im Deutschen und seinen Dialekten, *Zeitschrift für Dialektologie und Linguistik* LXXV, Franz Steiner Verlag, Stuttgart, pp. 282-331.

Perridon, H. 2003, Language contact and grammatical change: the case of Bergen. *Tijdschrift voor Skandinavistiek* 24, Skandinavisch Seminarium, Amsterdam, pp. 235-256.

Ralli, A. 2009, Morphology meets dialectology: insights from Modern Greek dialects. *Morphology* 19, pp. 87-105.

Refsum, H. 1954, *Romeriksmålet*, Romerike historielag, Oslo.

Sandøy, H. 2000, Vitnemål frå ein dativbrukar. Elmevik, L. (ed): *Dialekter och folkminnen. Hyllningsskrift till Maj Reinhammar den 17 maj 2000*. Sällskapet för svensk dialektologi, Uppsala, pp. 231-240.

Schøsler, L. 2001, From Latin to Modern French: Actualization and markedness. Andersen, H. (ed): *Actualization: Linguistic Change in Progress*. John Benjamins, Amsterdam, pp. 169-186.

Skjekkeland, M. 2005, *Dialektar i Noreg*, Høyskoleforlaget, Kristiansand.

Torp, A., Vikør, L.S. 2003, *Hovuddrag i norsk språkhistorie,* 3 edn, Ad Notam Gyldendal, Oslo.

Trask, R.L. 1996, *Historical Linguistics,* Arnold, London, 1996.

Trask, R.L. 2007, *Trask's Historical Linguistics*. 2nd edition, revised by Robert McColl Millar, Hodder Arnold, London.

Venås, K. 1993/1971, Eit oversyn over dei viktigaste endringar i språkstrukturen frå gammalnorsk til nynorsk Jahr, E.H. & O. Lorentz (eds), *Historisk språkvitenskap/Historical Linguistics*, Novus, Oslo, , pp. 253-266.

Wessén, E. 1969, *Svensk språkhistoria*. Almqvist &Wiksell, Stockholm.

Wetås, Å. 2008, *Kasusbortfallet i mellomnorsk*, Unipub, Oslo.

Wurzel, W.U. 1984, *Flexionsmorphologie und Natürlichkeit*, Berlin, Akademie-Verlag.

Wurzel, W.U. 2002, Nordic language history: Semiotics and the history of naturalness. Bandle, O. et al (eds), *The Nordic Languages, volume 1*, Mouton de Gruyter, Berlin, pp. 253-262.

Feature hierarchy and nominal inflection:
Evidence from ancient Greek

KONSTANTINOS KAKARIKOS
University of Athens
k_kakarikos@hotmail.com

1. Introduction

In many morphological theories and especially in those which consider morphology as an autonomous component of the grammar, features play a crucial role. They are considered to be linguistic primitives manipulated by the rules of word formation. However, despite their importance in both the description of language and the derivation of words, no serious attempts have been made so far with regard to their internal organization, namely the description of the internal structure of feature bundles. Previous works on the field such as Noyer's leading study (1992), the approaches of Harley (1994) and Harley and Ritter (2002) confine themselves to the study of the hierarchy of verbal features. Nominal features, although equally central in the description of the inflectional paradigm and the derivation of inflectional forms, have not been examined with regard to their interrelations under a general perspective.

In this paper we aim at the establishment of a nominal feature hierarchy for Ancient Greek on the basis of well-defined criteria and also at the examination of the possible relationship of this hierarchy to aspects of the Ancient Greek nominal inflection. More specifically, our proposal is first that the feature hierarchy is a language specific phenomenon founded on empirical facts and theoretical criteria and second that it is parallel to the role that individual features play in inflectional derivation and the distribution of the inflectional forms to a strictly defined number of inflectional patterns. Such an approach opens the way for a total reformation of our views on the derivation of the inflectional forms, the notion of the paradigm and the definition of features, like inflectional class. The theoretical framework of our approach is the Feature theory (Halle and Maranz 1994 and Ralli 1999, 2000, 2005 for Modern Greek) whereas a central point of the analysis is the view that morphology functions as an independent grammatical component whose principal role is the interpretation of syntactic information in its own terms; therefore it holds a post-syntactic position (Halle and Maranz 1993, 1994, Beard 1995, Stump 2006, Steward and Stump 2007, Sigurðsson 2009).

The term Ancient Greek concerns the Attic dialect of the classical period (5th – 4th centuries BC) whereas the linguistic data presented are all drawn from the grammatical descriptions of the dialect found in the literature (cf. Kühner and Blass 1890, Goodwin 1900 and especially Smyth 1976[10]).

The paper is organized as follows. Section 2 and its sub-sections express some general assumptions on Feature theory with a special focus on the type of nominal features and

their distribution with regard to the internal structure of nominal forms. Section 3 is a general introduction to the form of the nominal features of the Attic Greek. Section 4 deals with the definition of the theoretical criteria for the establishment of the hierarchy of nominal features for Attic Greek and also provides empirical evidence in support of this hierarchy. Section 5 examines the impact of the proposed feature hierarchy in the nominal inflection of the language, particularly in the structure of the operations involved in inflectional derivation and the organization of the inflectional patterns of the language, whereas section 6 presents some conclusive remarks.

2. Remarks on feature theory

Feature theory in Morphology has been developed on the model of feature organization in Phonology. According to this theory, morphological features are considered to be linguistic primitives with internal structure, exhibited in their organization into feature bundles. Each feature constitutes an abstract category which acquires a particular realization in certain linguistic environments. To capture this Feature theory organizes morphological features internally on an attribute – pair basis. Thus abstract features such as number, gender and case are the attributes while their particular realizations in certain environments like singular, plural, masculine, feminine, nominative, genitive etc are the values.

The definition of the morphological features of a particular language is a matter of parametric variation which allows for each language to select its own features from a large feature inventory in accordance to its internal morphological structure and organization. In morphologically rich languages feature selection as well as feature interrelation is realized through inflection. Each inflectional form corresponds to a number of features which codify the set of the grammatical information expressed by it. Therefore morphological features form the internal constitution of inflectional forms and trigger the morphological spelling operations which are responsible for inflectional derivation. The variety in feature selection attested cross-linguistically corroborates the differences of particular languages in their inflectional structure and their feature hierarchies. This remark is crucial for the definition of both the role and function of each feature in nominal inflection as it will be shown in the sections to follow.

2.1. The organization of the nominal features

In this section we put forward a number of general assumptions with regard to the organization of nominal features and the relation between this organization and the morphological processes involved in the derivation of inflectional forms. For many morphological theories (Halle and Vaux 1998, Ralli 2000, 2005), the internal constitution of each nominal form in terms of its abstract features is illustrated by the bipartite scheme given below in figure 1.

<center>Nominal Form</center>
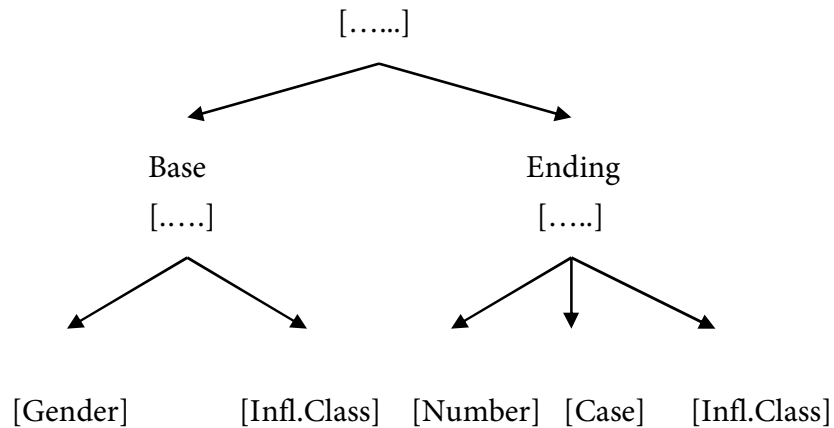
Figure 1

According to this scheme, each nominal form consists of a base and an ending. The base contains information about the grammatical category, the gender and the Inflectional Class (IC) of the form, while the ending carries information about the features of number and case. In some approaches (Ralli 1999, 2005) the feature of IC is also taken to be part of the ending in an attempt to secure its relation to a particular base or bases, a view which is also adopted here. In general terms, base features express information which contributes to the semantic identity of a form, whereas ending features are responsible for the morphological interpretation of abstract syntactic information.

However, this clear distinction of nominal features into semantic (lexical) and syntactic (functional) is often challenged by their type and behavior in both inflection and syntax. More specifically, the feature of gender, which generally expresses inherent semantic information and, for that reason, is considered as part of the base, participates also in agreement relations expressing syntactic relations as well (Ralli 1999, 2005). Quite similarly, the feature of number which is accommodated to the ending expresses also semantic information (Ralli 2005). Finally, the feature of IC which expresses no semantic or functional meaning and consequently belongs to both the base and the ending contradicts the binary distinction of the nominal features proposed (Ralli 2005, Sigurðsson 2009).

In view of these inconsistencies, we adopt a different model of analysis which is founded on the following assumptions:

(a)   Morphology is an autonomous component of the grammar of language whose principal role is to express syntactic information in its own terms. As such, it holds a position after syntax and its role is to interpret syntactic information (Halle and Maranz 1993, 1994, Beard 1995, Sigurðsson 2009).

(b)   Traditional morphemes are considered to be morphological spelling operations rather than formative elements. As such they have a functional load which is expressed in terms of features. Their function is to modify the phonological structure of the lexical bases over which they operate (Beard 1995).

(c)   The derivation of any inflectional form is a purely morphological process which is carried out by a set of morphological spelling operations in accordance to the information of the terminal node it fills which is defined by syntax and expressed in terms of syntactic features.

(d)   The strict relationship between syntax and morphology is to be understood on condition that there are two distinct feature sets, one syntactic and one morphological, which contain the same type of features and correspond to each other in a not always symmetric type of correspondence (Sigursson 2009).

Given this basis, we now proceed to a feature based analysis of the Attic Greek nominal inflection. Our aim is first to define the criteria for the establishment of a hierarchy for the nominal features of Attic Greek and second to investigate its impact (if any) to the inflectional derivation and the organization of inflection in general. Such an approach will possibly reform our view about the derivation of the inflectional forms as well as about the notions of the IC and the paradigm. However, before this, a short account of the nominal features of Attic Greek will precede.

3. The feature-based analysis of the Attic Greek inflection

The Attic Greek nominal inflection contains a set of four features each of which acquires a number of potential values as illustrated in (1):

(1)   Gender : Masculine (M), Feminine (F), Neuter (N)
          Number : Singular (Sg), Plural (Pl), Dual (Du)
          Case : Nominative (N), Genitive (G), Dative (D), Accusative (A), Vocative (V)
          Inflectional Class : 1 – 10 (IC)

From the abstract features gender, number and case are defined as arbitrary grammatical categories expressing lexical and/or syntactic information. IC on the other hand is generally considered as a conventional category which carries no lexical or functional load. Its role therefore is purely morphological, since it functions as a connective element defining the relationship between morphemes to their lexical bases (Ralli 2005, Sigurðsson 2009). For that reason it is always defined on the basis of specific criteria, which are either dictated by the theoretical model of analysis adopted or are imposed by the peculiarities of the morphological structure of the language under consideration. The criteria which have been proposed so far for the definition of the IC of standard Modern Greek and Ancient Greek are (i) the base allomorphy (Ralli 2000, 2005), (ii) the set of the endings attached to the various bases (Ralli 2000, 2005) and (iii) (especially for Ancient Greek) the stress pattern (Kakarikos 2009a, 2010). These criteria lead to the definition of a set of 10 values for the IC of Attic Greek which is presented in detail in the Appendix.

4. The hierarchy of nominal features: general assumptions

In general terms, the establishment of any feature hierarchy, in order to be methodologically well-organized, should be built on the basis of certain principles and conditions. First it should be defined as a language specific property (Stump 2006), second it should be able to accommodate both the abstract features as well as their special values and third it should be organized on the basis of theoretical criteria, having certain correspondences to the empirical facts of the language.

In our account, the first condition is satisfied by the fact that the proposed hierarchy of the nominal features refers exclusively to the Attic Greek dialect and therefore precludes generalizations of any sort; the second is satisfied by the setting of a general ranking for the abstract nominal features, like gender, number and case and also by the enrichment of this ranking with the addition of their special values, like masculine, feminine and neuter (for gender), singular, plural and dual (for number) and nominative, genitive, dative and accusative (for case)[10]. Finally, the third condition is satisfied first by the support of theoretical criteria, such as markedness, logical entailment and the like, which have been widely discussed in the literature on various occasions (cf. Noyer 1992 and Hurley 1994 among others) and second by the analogy to the typological and empirical facts of the dialect.

An important aspect of the analysis which should be stressed is that the proposed hierarchy does not include the feature of IC. This is due to its conventional character and the lack of any particular semantic or functional load related with it.

4.1. The hierarchy of the abstract features

A typical indicator of the relationship between the abstract nominal features, which have been pointed out in many approaches to inflection, is syncretism (cf. Baerman, Brown and Corbett 2005 and Stump 2006 among others). Syncretism is the morphological identification of forms which exhibit neutralization in some of their features. Both typological and empirical facts in many different languages show that in all cases of syncretism there is a fixed relation between the features neutralized and the features defining the context of neutralization. This sort of relationship is theoretically defined either as the *Dominance Hierarchy* (Hjelmslev 1935 cited by Carstairs 1987) or as the *Feature Ranking Principle* (Stump 2006) and states that the neutralized or dominated features are always ranked lower than the dominant features, namely the features defining the context of this neutralization.

This situation is compatible with the linguistic data of Attic Greek as illustrated in table 1.

Table 1: Types of syncretism in Attic Greek

	Features Neutralized			Morphosyntactic Context			Affixes Involved	Indicative Examples
	Gender	Number	Case	Gender	Number	Case		
1	-	-	N, A, V	N	Sg	-	/n/	*δῶρο-ν*

---

[10] Vocative is a special case with a different behaviour from other cases. As such it will not be discussed here.

2	-	-	N, A, V	N	Pl	-	/a/	δῶρ-α
3	-	-	N, A, V	M	Du	-	/e/	βότρυ-ε
4	-	-	G, D	F	Du	-	/in/	σκια-ὶν
5	-	Sg, Pl	D, N, V	M	-	-	/i/	λύκω-ι, λύκο-ι
6	-	Sg, Pl	G, A	F	-	-	/s/	σκιᾶ-ς

More specifically, in examples 1 – 4 the neutralized feature is case, whereas gender and number are the features determining the context of neutralization. Thus, according to the dominance hierarchy, gender and number dominate over case and therefore they are ranked higher as it is shown in (2).

(2)   GENDER, NUMBER > CASE

Similarly, in examples 5 – 6, the neutralized features are number and case, whereas the feature conditioning the neutralization is gender. Again, the dominance hierarchy ranks gender higher than number and case as it is shown in (3).

(3)   GENDER > NUMBER > CASE

The Feature hierarchy proposed in (3) finds further theoretical support by similar notions such as the *Relevance Hierarchy* (Bybee 1985, cited by Carstairs 1987) which classifies the nominal features according to their position with respect to the nominal base and their meaning. According to this, the more relevant features which are part of the base like gender are ordered higher than the less relevant features, like number and case, which are located on the right of the base. This sort of classification has certain correspondences to figure 1 above.

4.2. The hierarchy of the special feature values

Abstract feature ranking can be further analyzed and enriched with the addition of the special values which correspond to each of them. Special feature values are organized in feature sub-trees and are all subject to the control of their abstract counterparts. Their internal ranking is founded on three criteria:

(a)   The criterion of *markedness* which states that marked features are ranked lower than the unmarked ones.

(b)   The criterion of *logical entailment* (Noyer 1992, cited by Hurley 1994) which holds that features logically implied by other features of the same kind are ranked lower than them.

(c)   The criterion of *argument encoding* (Blake 2001) which refers particularly to the feature of case and states that case values encoding peripheral grammatical relations are ranked lower than case values encoding core relations.

All three criteria have certain applications to the hierarchy of the special feature values of Attic Greek.

More specifically, the criterion of *markedness* predicts the dominance of the animate (i.e. masculine and feminine) over the inanimate (i.e. neuter) gender[11], the dominance of the masculine over the feminine gender, the dominance of the singular number (considered as a means expressing individuality) over the plural and dual numbers (considered as means expressing aggregation), and finally the dominance of the direct over the indirect cases.

The criterion of *logical entailment* specifies the relationship of the plural and dual by predicting the predominance of the plural in the sense that duality is a notion already existing in the expression of aggregation.

Finally, the criterion of *argument encoding* sheds light to the internal organization of the special values of the case feature in a way reflecting the distinction of the grammatical relations into core (i.e. subject < direct object < indirect object) and peripheral (i.e. oblique meanings, such as locative, instrumental etc).

The discussion on the hierarchy of the special feature values leads to a more elaborated version of (3) which is given in figure 2.
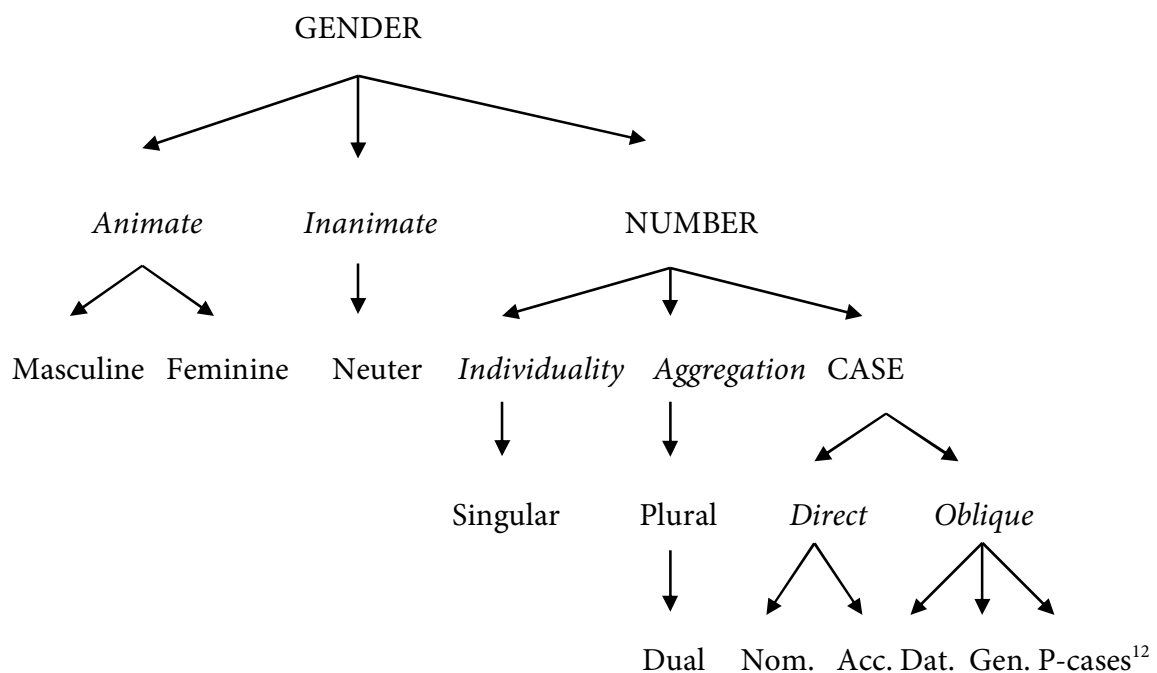
Figure 2

The hierarchy proposed in figure 2 finds empirical support in the linguistic data of Attic Greek. Thus, the internal hierarchy of gender values is reflected in agreement

---

[11] The ranking of animate > inanimate is founded on typological evidence from many languages (Corbett 1999).

[12] P-cases concern structures with a preposition and a noun and are not to be discussed here.

constructions defined according to the scheme: human entities > non-human entities > inanimate entities (cf. Chila-Markopoulou 2003). Similarly, the internal relation of the plural and dual is reflected in cases of verbal agreement where the verb in the plural agrees with the subject in the dual and vice versa. Consider the examples in (4) and (5) below.

(4) *Ξενοφῶντι προσέτρεχον* (PL) *δύω*(DU) *νεανίσκω*(DU) (*Xen. Anab.4.3.19*)

"Two youths ran up to Xenophon"

(5) *Δύο ἄνδρες*(PL) *προσελθόντε*(DU) *Ἄγιδι διελεγέσθην*(DU) *μή ποιεῖν μάχην* (*Thuc.5.59.5*)

"Two men coming to Agis urged him not to fight"

Finally, the hierarchy of the case feature values is reflected in the numerous instances of syntactic conflict which are exemplified by case attraction phenomena and resolved on the basis of case feature hierarchies (Grosu 1994, Vogel 2001, 2003 and Kakarikos 2009b). Consider the examples in (6), (7), (8) and (9) (Kakarikos 2009b).

Accusative vs. Nominative → Accusative

(6) *τοῦ ναυτικοῦ αὐτῶν* **τό πλῆθος ὅ** *πρίν οὐχ ὑπῆρχε* (*Thuc.1.90*) (instead of *ὅ πρίν*)

"the greatness of their navy, which there was not before"

Dative vs. Accusative → Dative

(7) *τοῖς γάρ* **ἀγαθοῖς οἷς** *ἔχομεν ἐν τῇ ψυχῇ* (*Isocr.8.32*) (instead of *ἀγαθοῖς ἅ*)

"with the goodness which we have in our soul"

Genitive vs. Dative → Genitive

(8) *τινές* **ὀλίγοι ὧν** *ἐγώ ἐντετύχηκα* (*Pl.Res.531e*) (instead of *ὀλίγοι οἷς*)

"a few of those whom I have met with"

P-case vs. Dative → P-case

(9) *σύν ταῖς δώδεκα* **ναυσίν αἷς** *εἶχεν* (*Xen.Hell.4.8.23*) (instead of *ναυσίν ἅς*)

"with the twelve ships which he had"

4.3. Interim summary: remarks on feature hierarchy

A close examination of the feature hierarchy proposed in figure 2 is rather informative as regards the definition of the principles underlying the interrelations of the various nominal features inside the sphere of the hierarchy. These are:

(a)   the *Dominance Principle*: higher ranked features dominate all lower ranked features.

(b)   the *Dependency Principle*: higher ranked features have less dependents than lower ranked features.

(c)   the *Complexity Principle*: higher ranked features are less complex than lower ranked features, since they are dominated by fewer feature nodes.

Thus, according to the *dominance principle* the feature of gender dominates over the features of number and case on the basis of its higher ranking. In the same vein, the feature of number dominates over the case feature. According to the *dependency principle*, gender and number as higher ranked features, have three dependents (i.e. masculine, feminine, neuter and singular, plural, dual, respectively) whereas case, as a lower ranked feature, has five (i.e. nominative, genitive, dative, accusative and the p-cases). Finally, according to the *complexity principle*, case as the feature filling the lowest position in the hierarchy and therefore dominated by two nodes (i.e. gender and number) is more complex than number which is dominated by only one node (i.e. gender) and gender which has no dependence on any other feature node.

The establishment of a well-defined hierarchy for the nominal features of Attic Greek leads to the second aim of our analysis which concerns its impact to both the role of nominal features in inflectional derivation processes and the definition of the inflectional patterns allowed in that dialect. More specifically, the discussions concern first the possible analogy between the ranking of each nominal feature proposed in figure 2 and its role in triggering the morphological spelling operations for the formation of the inflectional forms and second the possible connection of the feature hierarchy with the (fixed) number of the inflectional patterns of Attic Greek. This second issue reforms our view for both the notion and role of the inflectional paradigm in the sense it is considered as an entity not existing in advance as a well-defined and fixed set of inflectional forms but rather as an entity arising from the derivational mechanism of morphology.

5. Feature hierarchy and nominal inflection

In a theoretical model which accepts a post-syntactic position for morphology, inflectional derivation is carried out by morphological operations which are triggered by syntax. Given this, the main aim of morphology is to derive lexical forms for the terminal positions of a syntactic structure. A necessary condition for this is the existence of a correspondence between the information underlying the terminal positions which is syntactic and the information expressed by the inflectional forms, which is morphological. Such an approach presupposes the existence of two feature sets, one syntactic and one morphological. The correspondences between the two feature sets are not always symmetric and this may explain the possible asymmetries often existing between syntax and morphology (cf. Ralli 2005). On the other hand, it reforms the view which classifies nominal features into lexical and syntactic, by considering them all as morphological entities which realize abstract syntactic information in their own terms (Sigurðsson 2009; for a similar view cf. also. Booij 1995 and Ralli 2005). In such a context, morphological features take different roles which are reflected both in the

inflectional derivation processes and the definition of the inflectional patterns of a language. These roles will be further analyzed and described in the following sections. A crucial point of the discussion is that it supports the autonomy of morphology from a different point of view which does not dismiss its interface with syntax.

5.1. Feature hierarchy and inflectional derivation processes

The importance of the role of morphological features and their hierarchical ranking in inflectional derivation may be found in connection with the morphological spelling operations which are involved in the derivation of the inflectional forms of Attic Greek and presented in table 2 along with their feature specifications.

Table 2. The morphological spelling operations of Attic Greek

MORPHOLOGICAL SPELLING OPERATIONS		NOMINAL FEATURE SPECIFICATIONS			
		GENDER	NUMBER	CASE	IC
1	/s/	{M, F}	{Sg, Pl}	{Nom, Gen, Acc, Voc}	{1 – 6, 8, 10}
2	/o/	{M, F, N}	Sg	Gen	{1 – 4}
3	/i/	{M, F, N}	{Sg, Pl}	{Nom, Dat, Voc}	{1 – 10}
4	/n/	{M, F, N}	Sg	{Nom, Acc, Voc}	{1 – 6}
5	/ø/	{M, F, N}	Sg	{Nom, Acc, Voc}	{1 – 10}
6	/ɔːn/	{M, F, N}	Pl	Gen	{1 – 10}
7	/is/	{M, F, N}	Pl	Dat	{1 – 4}
8	/si/	{M, F, N}	Pl	Dat	{5 – 10}
9	/e/	{M, F, N}	Du	{Nom, Acc, Voc}	{1 – 10}
10	/in/	{M, F, N}	Du	{Gen, Dat}	{1 – 4}
11	/oin/	{M, F, N}	Du	{Gen, Dat}	{5 – 10}
12	/os/	{M, F, N}	Sg	Gen	{5 – 10}
13	/es/	{M, F}	Pl	{Nom, Voc}	{5 – 10}
14	/as/	{M, F}	Pl	Acc	{7 – 10}
15	/a/	{M, F, N}	{Sg, Pl}	{Nom, Acc, Voc}	{2 – 10}

The first point which emerges from table 2 is that all morphological operations are always triggered by a combination of features; therefore no feature has the complete control over an inflectional derivation process. However, such a view does not preclude the existence of an internal scale in the role of particular features inside the general frame of the feature specifications for each operation. The main reason for this is that some features are more flexible than others in distinguishing morphological operations by specifying them more precisely. This flexibility is consistent with the number of their special values as well as their position in feature hierarchy. Thus, complex features with many special values have a greater control to the inflectional derivation processes than less complex features. This is because they are able to define the maximum number of morphological operations which could apply to the lexical bases of a language and accordingly the maximum number of the correct inflectional formations possible in it.

For instance, according to table 2 the possible feature combinations for the morphological spelling operation suffixing /s/ can theoretically lead to the derivation of 16 inflectional forms ending in /s/ (i.e. 2 x 2 x 4 = 16). Since this is not true for the Attic Greek inflection, the blocking of the ungrammatical formations is related to the role that each morphological feature plays in any inflectional derivation process. More concretely, the specification of the morphological operations suffixing /s/ by the morphological feature of gender alone is quite general and vague. This situation is improved with the addition of the feature of number. The combination of the features of gender and number make more clear predictions as to the functional load of the derived forms and also increases their number from 2 to 4 (i.e. 2 x 2 = 4: 1. [M, Sg], 2. [M, Pl], 3. [F, Sg], 4. [F, Pl]). All forms are taken to be grammatical. The whole situation becomes even more precise with the addition of the feature of case.

Case has the greatest number of special values and as a result it increases the number of possible formations from 4 to 16 (i.e. 4 x 4 = 16: 1. [M, Sg, N], 2. [M, Sg, G], 3. [M, Sg, A], 4. [M, Sg, V], 5. [M, Pl, N], 6. [M, Pl, G], 7. [M, Pl, A], 8. [M, Pl, V], 9. [F, Sg, N], 10. [F, Sg, G], 11. [F, Sg, A], 12. [F, Sg, V], 13. [F, Pl, N], 14. [F, Pl G], 15. [F, Pl A], 16. [F, Pl, V]). However from this set of candidate forms, only 9 are grammatical. The blocking of the ungrammatical formations is controlled to some extend by the feature of case as well as by the feature of IC.

To understand how this is possible, we focus on the formation of a subset of the forms ending in /s/, namely the forms derived by operations specified as [M, Pl] and [F, Pl]. Both operations, although quite general as far as their specification is concerned, are theoretically considered to lead to grammatical formations. This situation radically changes with the addition of the feature of case. Case feature has the greatest set of feature values (i.e. Nominative, Genitive, Accusative, Vocative) and therefore may define morphological operations in a more precise way. According to this, from the possible 8 candidate forms which come out from the combination of the features of gender (M, F), number (Sg, Pl) and case (N, G, A and V), (i.e. 1. [M, Pl, N], 2. [M, Pl, G], 3. [M, Pl, Acc], 4. [M, Pl, V], 5. [F, Pl, N], 6. [F, Pl, G], 7. [F, Pl, A], 8. [F, Pl, V]) only two are grammatical (i.e. the combinations: [M, Pl, A] and [F, Pl, A]). All other forms are blocked as ungrammatical. This blocking is the combined result of the coordination between the features of case and IC. The role of case concerns the precise and analytic definition of the candidate forms bearing the suffix /s/ which may be theoretically derived. The role of the IC concerns the filtering of their feature specifications leading to the emergence of the correct formations.

The example is very illustrative since it gives an idea about the number of forms which are finally derived and take their place in inflection. Moreover it explains why from a large number of possible combinations grammar allows the derivation of a small set of forms. In this context it supports the role of the feature of case as the leading feature in the hierarchy of features compared to both gender and number. On the other hand, it does not dismiss the role of IC as an equally central feature in inflectional derivation despite its absence from the hierarchy of nominal features and highlights its character as

a purely morphological element serving formative aims with no reference to syntax and semantics.

Summing up all three features of gender, number and case contribute to the derivation of the right inflectional form; however their contribution varies depending on the number of their special values. As a result, the number of the grammatical inflectional forms in a language is gradually limited as the analytic definition of the morphological spelling operations with the addition of more features is increased. And the role of features in this procedure is attuned to the feature hierarchy proposed in figure 2.

5.2. Feature hierarchy and the inflectional patterns of language

The administration of the inflectional formation processes in the way described above plays also a decisive role in the organization of the derived inflectional forms into inflectional patterns. More specifically, a common feature of the fusional languages is the organization of the inflectional forms into larger groups, traditionally known as paradigms. Each paradigm exhibits its own combination of endings and represents a special and unique pattern of inflection. The definition of the inflectional patterns of a language depends on its morphological structure and is carried out with the most economical way possible for that language. Therefore, the number of the inflectional patterns of a language is always fixed in a way which does not allow for changes of any sort. In Attic Greek of the classical period (5th – 4th centuries BC) the number of the inflectional patterns is equal to 8, as it is shown in table 3, below.

Table 3. The inflectional patterns of Attic Greek (Kakarikos 2007, 2010)

PATTERN	I	II	III	IV	V	VI	VII	VIII
IC	1, 3 – 4	1 – 3	5 – 6	8	10	7, 9	3 – 4	5 – 9
GENDER	M / F	F	M / F	M / F	M	M / F	N	N
SINGULAR								
NOM.	-s	-ø	-s	-s	-s	-ø	-n	-ø
GEN.	-o	-s	-os	-os	-os	-os	-o	-os
DAT.	-i	-i	-i	-i	-i	-i	-i	-i
ACC.	-n	-n	-n	-a	-a	-a	-n	-ø
VOC.	-ø	-ø	-ø	-s	-ø	-ø	-n	-ø
PLURAL								
NOM.	-i	-i	-es	-es	-es	-es	-a	-a
GEN.	-ɔːn	-ɔːn	-ɔːn	-ɔːn	-ɔːn	-ɔːn	-ɔːn	-ɔːn
DAT.	-is	-is	-si	-si	-si	-si	-is	-si
ACC.	-s	-s	-s	-as	-as	-as	-a	a
VOC.	-i	-i	-es	-es	-es	-es	-a	a
DUAL								
NOM.	-e	-e	-e	-e	-e	-e	-e	-e
GEN.	-in	-in	-oin	-oin	-oin	-oin	-in	-oin
DAT.	-in	-in	-oin	-oin	-oin	-oin	-in	-oin
ACC.	-e	-e	-e	-e	-e	-e	-e	-e
VOC.	-e	-e	-e	-e	-e	-e	-e	-e

This situation raises questions regarding the factors conditioning the formation of the inflectional patterns in a way which allows for the emergence of the right combinations and blocks the wrong ones. The possible answer to these questions is related with the role of features in controlling both the systematic differences and similarities of each

pattern, a function which once more is strictly connected to their special position in the feature hierarchy.

The starting point of the discussion is that the formation of the inflectional patterns of a language depends on the inflectional derivation processes. As a result, inflectional patterns may be better understood as units which emerge via a number of morphological operations and therefore do not exist *a priori*. Such a view is against the traditional notion of paradigm, but is consistent with our view of morphology as an autonomous post-syntactic grammatical component which derives forms on the basis of syntactically driven information.

On the other hand, as it comes out from the empirical facts of Attic Greek, the idea for the *a priori* existence of the paradigm gives only a partial explanation both to its internal organization as well as to the number of paradigms which are possible in that dialect. This is because the treatment of nominal features outside the derivational processes which give rise to the inflectional forms puts limits to their role as factors conditioning inflection.

In such a context, gender seems to capture the systematic differences across different patterns but only in the range of certain sub-systems. is the case with the differences in the nominative and genitive singular of the inflectional patterns I and II which is controlled by the distinctions between the masculine and feminine genders (e.g. (M) *neaniaː-s* (νεανίας) vs. (F) *moːsa-∅* (μοῦσα), (M) **neanio-o → neanioː-∅* (νεανίου) vs. (F) *moːseː-s* (μούσης)) or the difference between the inflectional patterns I, III, IV, V and the inflectional pattern VII as well as that between the inflectional patterns I, III, V and VIII in the nominative and vocative singular respectively, which is controlled by the distinction between the masculine and feminine genders from the neuter (e.g. (M/F): *lyko-s/neːso-s* (λύκος/ νῆσος) vs. (N): *dɔːro-n* (δῶρον), (M/F): *botry-s/ophry-s* (βότρυς/ οφρῦς) vs. (N): *dakry-∅* (δάκρυ) (M/F): *peleky-s/ poli-s* (πέλεκυς/ πόλις) vs. (N): *asty-∅* (ἄστυ) (M/F): *korak-s/lamba-s* (κόραξ/ λαμπάς) vs. (N): *sɔːma-∅* (σῶμα) etc).

Quite similarly, the feature of number captures the differences across all inflectional patterns but again in the range of its immediate control, namely the five cases of each number, while it is unable to capture both differences and similarities across different numbers and paradigms (e.g. (Sg): *lykos* (λύκος), **lyko-o → lykoː* (λύκου), *lykɔː-i* (λύκω) etc vs. (Pl): *lyko-i* (λύκοι), **lyko-ɔːn → lyk-ɔːn* (λύκων), *lyko-is* (λύκοις) etc).

In contrast to both number and gender, case seems to be more flexible. However it lacks the basis which could define it as the feature connecting the various forms inside a paradigm and also distinguishing them from the forms of other paradigms.

A different and more thorough approach could come through the examination of the inflectional patterns as units emerging through a number of derivational processes which, apart from conditioning the derivation of the right inflectional formations, are also responsible for their classification into the patterns of inflection.

This sort of classification is envisaged on the basis of a vertical and a horizontal axis. The vertical axis defines the role of features in controlling the cohesion of forms belonging to the same inflectional pattern. The horizontal axis concerns the role of

features in defining the relations between forms belonging to different inflectional patterns. In the first case what matters is the control over the similarities in the feature specification of each inflectional form. In the second case the interest focuses on the control over the differences in the feature specification of each inflectional form.

A necessary prerequisite for the first task is the variation in the range of the special values of the controlling feature. The more special values a feature has, the most powerful as a factor conditioning similarities (that is to say grouping similar forms) it is. For instance, the feature of gender fails to control the similarity of forms such as *anthrɔːpo-s*, *ˣanthrɔːpo-o* ➔ *anthrɔːpoː* … etc of the Inflectional Pattern I, since it is the only variable available in this pattern. On the other hand, the feature of number which has a greater range of variables is able to control more similarities. Consider the forms *anthrɔːpo-s*, *ˣanthrɔːpo-o* ➔ *anthrɔːpoː* etc. (which are marked as Masculine, Singular) vs. the forms *anthrɔːpo-i*, *anthrɔːp-ɔːn* etc (which are all marked as Masculine, Plural) and the forms *ˣanthrɔːpo-e* ➔ *anthrɔːpoː*, *anthrɔːpo-in* etc (which are marked as Masculine, Dual). However each value of the number feature fails to control the similarities of forms inside its range of control (cf. the forms *anthrɔːpo-s*, *ˣanthrɔːpo-o* ➔ *anthrɔːpoː* etc) which are specified as Masculine, Singular). Therefore its controlling power is greater than genders but it is still limited compared with that of case. This is because case has an even greater number of values than both gender and number and therefore a greater range of control over the features of the same inflectional pattern. Thus the similarities of the forms *anthrɔːpo-s*, *ˣanthrɔːpo-o* ➔ *anthrɔːpoː*, *anthrɔːpoː-i*, *anthrɔːpo-n*, *anthrɔːpe-∅* are controlled by the feature of case which again is the only varying feature in the feature specification of all five forms (cf. M, Sg, **N** vs. M, Sg, **G** vs. M, Sg, **D** vs. M, Sg, **A** vs. M, Sg, **V**, respectively) (for a similar view cf. Blevins 2006).

Summing up, features with a great number of special values, such as case, play a more decisive role in controlling the cohesion inside the same pattern of inflection as opposed to features with less special values, such as number. This distinction is consistent with the proposed feature hierarchy.

The second task, namely the control over the differences of forms belonging to different patterns of inflection, is a consequence of the first. It depends on the number of the different forms which may be grouped under the heading of a single feature in the most precise way possible. This possibility is however related to the power of different features in controlling the cohesion inside the same pattern. Thus features with few special values and small power in controlling the cohesion of an inflectional pattern, allow for large groupings which are vague and unclear. In contrast, features with more special values, have a wider range of choices and therefore are more powerful in controlling the cohesion of the forms inside the same pattern. As a result they allow for more concrete and precise classifications by managing a more strict and solid organization of inflection in general.

In this context the grouping of the inflectional endings of Attic Greek according to the feature of gender alone, gives 15 possible inflectional patterns (on the basis of the feature with the greatest number of endings). Similarly, a grouping on the basis of gender

reveals a number of 9 possible inflectional patterns. In contrast to both gender and number, the feature of case allows for a number of 7 patterns of inflection. Consider table 4 below.

Table 4. Possible definitions of the inflectional patterns of Attic Greek

(a) Definition according to gender alone

Allomorphs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Masculine	s	o	i	n	∅	ɔːn	is	e	in	os	es	si	as	oin	a
Feminine	∅	s	o	i	n	ɔːn	is	e	in	os	es	si	oin	a	as
Neuter	n	o	i	a	ɔːn	is	e	in	∅	os	si	oin	-	-	-

(b) Definition according to number alone

Allomorphs	1	2	3	4	5	6	7	8	9
Singular	s	∅	n	o	os	i	a	-	-
Plural	i	es	a	ɔːn	is	si	s	as	a
Dual	e	in	oin	-	-	-	-	-	-

(c) Definition according to case alone

Allomorphs	1	2	3	4	5	6	7
Nom	s	∅	n	i	es	a	e
Gen	o	s	os	ɔːn	in	oin	-
Dat	i	is	si	in	oin	-	-
Acc	n	a	∅	s	as	in	oin
Voc	∅	s	n	i	es	a	e

As we see in the table 4, there is again an internal scale as regards the role of each feature in the organization of the inflectional patterns which is in absolute accord with the feature hierarchy proposed in figure 2. More specifically, complex features, like case, allow for more precise classifications as opposed to less complex features, like gender and number. This hierarchical difference may be better seen in the way that all inflectional features coordinate for the formation of an inflectional pattern. For instance, in a possible formation of the inflectional patterns of, say, the Masculine Singular, the 15 allomorphs defined by the gender feature alone (cf. table 4a) are reduced to 7 with the addition of the feature of number (cf. table 4b). This number of 7 allomorphs is increased to 15 with the addition of the feature of case (cf. table 4c). However, this is an occasional increase since this number (except from the endings which are defined as plural or dual, e.g. -es, -is, -si, -in etc) includes also ill-formations (i.e. the definition of /*i/, /*n/, /*a/ as [M, Sg, N], the definition of /*s/ as [M, Sg, G] and [M, Sg, A] and finally the definition of /*i/, /*a/, /*s/ and /*n/ as [M, Sg, V]) which are blocked as ungrammatical. Therefore the final number is fixed to 7 endings (i.e. /s/, /Ø/,

/o/, /os/, /i/, /n/, /a/). Interestingly, this number of 7 endings does not point to 7 inflectional patterns, since their distribution is controlled (apart from the case feature) by other factors, such as the feature of IC and the influence of the rules of syncretism. IC distributes similar endings into different inflectional patterns (e.g. /s/: IC 1, 3 – 6, 8, 10, /os/: IC 5 – 10 etc. Cf. table 3) while syncretism is responsible for the distribution of similar endings into the same inflectional pattern (e.g. /s/: inflectional pattern 1 – 4, /os/: inflectional pattern 2 – 4 etc. Cf. table 3).

This example corroborates the hierarchical role of each nominal feature in inflection. However it cannot predict the exact organization of forms to the vertical axis. This means that, although case may secure the internal cohesion of the inflectional form inside an inflectional pattern, the formation of this pattern, namely the exact combinations of forms in a certain way and not in another is again a matter strictly related to the feature of IC. This is a further indication which stresses its importance as a purely morphological formative element, despite the lack of any semantic or functional load. Thus, according to table 3, the 7 endings of our example will be distributed in a fixed set of 5 patterns defined by the feature of IC.

A final remark from this discussion which should be pointed is the reformation of our view about the type, organization and function of the traditional paradigm. This is no more defined as a well-organized set of forms which exist in advance and to which all inflectional derivation processes refer, but rather as a dynamic grammatical entity which emerges through morphology as a result of the coordination of nominal features in the general framework of inflectional derivation.

6. Conclusions

Summing up, in this paper we argued that nominal features may be organized into language specific hierarchies. These hierarchies are organized on the basis of empirical (i.e. syncretism) and theoretical well-defined criteria (i.e. markedness, logical entailment and argument encoding) and also have certain implications to the role of individual features in both the inflectional derivation processes and the organization of the inflectional patterns of a language.

More specifically, less complex features, such as gender and number, have a limited control over inflection whereas more complex features, such as case, have a more complicated and decisive role. In practical terms, this hierarchy is expressed by the dominance of case over the other features and the use of the information provided by them first to the formation of the inflectional forms of the language by triggering the relevant morphological spelling operations and second to the distribution of the derived forms into a fixed set of inflectional patterns.

However, it should be stressed that this hierarchical ranking is built on the assumption that nominal inflection is the combined result of a set of features and therefore it does not prove in any way the absolute dominance of particular feature over it. On the other hand, features, such as the IC, although excluded from the feature hierarchy because of their conventional character, have still a crucial role as formative elements, purely

morphological in nature, which form the reference point of all derivational processes and support the role of the other features.

Practically speaking, the establishment of a feature hierarchy in inflection reveals many aspects of the function of both the operations involved in the derivation of the inflectional forms of a language and their organization into patterns of inflection. Moreover it improves our view on the notion of the inflectional paradigm by defining it as a derived, rather than as a pre-existing set of forms which emerges by means of the deductive logic via morphological operations built on morphological features and reflecting their hierarchical status.

Appendix

The inflectional classes of Attic Greek (Kakarikos 2007, 2009a)

IC	GENDER	STEM ALLOMORPHY	STEM VARIABLES	GREEK FORMS	GLOSSES
1	M	-aː ~ -a	neaniaː- ~ neania-	νεανίας	"young man"
		-ɛː ~ -a	poiɛːtɛː- ~ poiɛːta-	ποιητὴς	"poet"
	F	-aː ~ -a	skiaː- ~ skia-	σκιά	"shadow"
		-ɛː ~ -a	korɛː- ~ kora-	κόρη	"girl"
2	F	-a ~ -aː	alɛːthejja- ~ alɛːthejjaː-	ἀλήθεια	"truth"
			muːsa- ~ muːsaː-	μοῦσα	"muse"
3	M	-a	bo̥rrea-	Βορρᾶς	"North"
			hermea-	Ἑρμῆς	"Hermes"
		-o	ploo-	πλοῦς	"sailing"
	F	-a	mnaa-	μνᾶ	"mna"
		-a	sykea-	συκῆ	"fig tree"
	N	-o	osteo-	ὀστοῦν	"bone"
4	M / F	-o	lyko-, nɛːso-	λύκος, νῆσος	"wolf", "island"
		-ɔː	leɔː-, halɔː-	λεώς, ἅλως	"people", "thrashing floor"
	N	-o	dɔːro-	δῶρον	"present"
		-ɔː	anɔːgeɔː-	ἀνώγεων	"upper floor"
5	M / F	-iː	kiː-	κίς	"worm"
		-y	botry-	βότρυς	"grapes"
	N	-y	dakry-	δάκρυ	"tear"
6	M / F	-yː ~ -y	ophryː- ~ ophry-	ὀφρῦς	"eyebrow"
		-i ~ -e	poli- ~ pole-	πόλις	"city"
		-y ~ -e	peleky- ~ peleke-	πέλεκυς	"axe"
		-ɔːw ~ -ow	bɔːw- ~ bow-	βούς	"ox"
		-aːw ~ -aw	graːw ~ -graw, naːw ~ -naw	γραῦς, ναῦς	"old woman", "ship"
	N	-y ~ -e	asty- ~ aste-	ἄστυ	"city"
		-es	Sɔːkrates-, Periklees-, triɛːres-	Σωκράτης, Περικλῆς, τριήρης	"Sokrates", "Perikles", "trireme"
		-Vnt	geront-	γέρων	"old man"

7	M / F	-Vn	daimon-	δαίμων	"god"
		-Vr	r̥ɛ:tor	ῥήτωρ	"orator"
			pater-, mɛ:ter-	πατὴρ, μήτηρ	"father", "mother"
		-os	aidos-	αἰδῶς	"shame"
		-o	hekho-	ἠχῶ	"echo"
	N	-es	beles-, khrees-	βέλος, χρέος	"arrow", "dept"
		-as	kreas-	κρέας	"meat"
8	M / F	-C	korak-, pteryg-, phleb-, tapɛ:t-, lambad-, patrid-, akti:n-	κόραξ, πτέρυξ, φλὲψ, τάπης, λαμπὰς, πατρὶς, ἀκτὶς	"raven", "wing", "vein" "carpet", "torch", "country", "ray"
		-Vnt	imant-, odont-	ἱμὰς, ὀδοὺς	"strap", "tooth"
	N	-t	sɔ:mat-	σῶμα	"body"
9	M/F	-Vn	poimen-	ποιμὴν	"shepherd",
		-Vr	aither-	αἰθὴρ	"air",
		-V:n	hellɛ:n-, kheimɔ:n -	Ἕλλην, χειμὼν	"Greek", "winter"
		-V:r	klɛ:tɛ:r-	κλητὴρ	"usher"
	N	-r	nektar-	νέκταρ	"nectar"
10	M	-ɛ:w	basilɛ:w-	βασιλεὺς	"king"
		-Vnt	gigant-	γίγας	"giant"

References

Beard, R. 1995. *Lexeme - Morpheme Base Morphology. A general Theory of Inflection and Word Formation.* SUNY Series in Linguistics. Albany: State University of New York Press.

Baerman M., D. Brown and G. Corbett. 2005. *The Syntax – Morphology Interface. A study of Syncretism.* Cambridge: Cambridge University Press.

Blake, B. 2001². *Case.* Cambridge: Cambridge University Press.

Blevins, J. 2006. Case and Declensional Paradigms. In A. Malchucov and A. Spencer (eds) *The Oxford Handbook of Case*, 200 – 218. Oxford: Oxford University Press.

Booij, G. 1995. Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij and J. von Marle (eds). *Yearbook of Morphology 1995*, 1 – 16. Dortrecht: Kluwer.

Bybee, J. 1985. *Morphology. A study of the Relation between Meaning and Form.* Amsterdam. John Benjamins.

Chila-Markopoulou, D. 2003. Gender and Agreement in Modern Greek [In Greek]. In A. Anastasiadi-Simeonidi, A. Ralli and D. Chila-Markopoulou (eds). *Gender*, 132 – 167. Athens: Patakis.

Carstairs, A. 1987. *Allomorphy in Inflection.* Croom Helm Linguistics Series. London: Croom Helm.

Corbett, G. 1999. *Gender.* Cambridge: Cambridge University Press.

Goodwin, W.W. 1900. *A Greek Grammar.* Boston: Ginn and Company.

Grosu, A. 1994. *Three studies in Locality and Case.* London: Routledge.

Halle and A. Maranz 1993. Distributed Morphology and Pieces of Inflection. In *The view from Building 20*: Essays in Honor of Sylvain Bromberger. Edited by Kenneth Hale and Samuel Jay Keyser. Cambridge MA: MIT Press, 111 – 176.

Halle and A. Maranz 1994. Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics* 21, 275 – 288.

Halle and B. Vaux 1998. Theoretical Aspects of Indo–European Nominal Morphology: The Nominal Declensions of Latin and Armenian. In J. Jasanoff et al. (eds). *Mír Curad. Studies in Honor of Calvert Watckins*. Innsbruck: Innsbrucker Beiträge zur Sprachenwissenschaft, 223 – 240.

Harley, H. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. In Andrew Carnie and Heidi Harley (eds). *Papers on phonology and morphology: MIT Working Papers in Linguistics* 21, 289 – 320. Cambridge, MA: MIT Working Papers in Linguistics, Department of Linguistics, MIT.

Harley, H. and E. Ritter 2002. Structuring the Bundle: A Universal Morphosyntactic Feature Geometry. In H. Simons and H. Weise (eds). *Pronouns: Features and Representation*, 23 – 39. Amsterdam: John Benjamins.

Hjelmslev, L. 1935. *La catégorie des cas, étude de grammaire générale*. Acta Jutlandica, VII, 1.

Kakarikos, K. 2009a. *Morphological and "Semantic" Examination of the Case System of Ancient Greek* [In Greek]. PhD Thesis. Athens: University of Athens.

Kakarikos, K. 2009b. Case Conflicts in Ancient Greek Relative Constructions. A Feature-based Approach. Paper presented at the *XIXth International Conference on Historical Linguistics*. Nijmegen, 10 – 15 August 2009.

Kakarikos, K. 2010. A featured-based analysis of the Attic Greek Nominal Inflection. To appear in I. Putzu, G. Paulis, G. Nieddu and P. Cuzzolin (eds). *La morfologia del greco tra tipologia e diacronia. Atti del VII Incontro internazionale di linguistica greca* (Cagliari 13-15 settembre 2007), 239-255. Milano: Franco Angeli.

Kühner, R. and F. Blass 1890. *Ausführliche Grammatik der Griechischen Sprache. vol. I.* Hannover: Hansche Buchhandlung.

Noyer, R. 1992. *Features, Positions and Affixes in Autonomous Morphological Structure*. MIT PhD dissertation. Cambridge: MIT Working Papers in Linguistics.

Ralli, A. 1999. Inflectional Features and the Morphological Module Hypothesis. In Ath. Kakouriotis and V. Bolla-Mavrides (eds). *Working Papers in Theoretical and Applied Linguistics*, 6, 111 – 141. Thessaloniki.

Ralli, A. 2000. A Feature-based Analysis of Greek Nominal Inflection. *Glossologia* 11-12, 201 – 228.

Ralli, A. 2005. *Morphology* [In Greek]. Athens: Patakis.

Sigurðsson, H. 2009. Remarks on Features. In K. Grohmann (ed). *Explorations of Phase Theory: Features and Arguments*, 21 – 52. Berlin: Mouton de Gruyter.

Smyth, H. W. 1976[10]. *Greek Grammar*. Cambridge Mass: Harvard University Press.

Steward, Th. and H. Stump 2007. *Paradigm Function Morphology and the Morphology – Syntax Interface.* In Ramchand, G. and Ch. Reiss (eds). *The Oxford handbook of linguistic Interfaces*, 383 – 420. Oxford: Oxford University Press.

Stump, G. 2006. *Inflectional Morphology*. A Theory of Paradigm Structure. Cambridge: Cambridge University Press.

Vogel, R. 2001. Case Conflict in German Free Relative Constructions. An Optimality Theoretic Treatment. In G. Müller and W. Sternefeld (eds). *Competition in Syntax*. 341 – 375. Berlin: Mouton de Gruyter.

Vogel, R. 2003. Surface Matters. Case conflict in the free relative constructions and case theory. In E. Bradner and H. Zinsmeister (eds). *New perspectives in Case Theory*. 269 – 300. Standford: CSLI Publications.

Locative alternation and verb compounding in Japanese[1]

HIDEKI KISHIMOTO
Kobe University
kishimot@lit.kobe-u.ac.jp

0. Abstract

Locative alternation verbs can invoke an alternation in the morphological frames of arguments, expressing the meanings of motion and change of state. In this paper, on the basis of the alternation patterns of compound verbs in Japanese, it is shown that locative alternation verbs are divided into two classes. One class of verbs is formed by what Pinker defines as a 'perspective shift', involving a derivation from one variant to the other. Verbs in the other class are inherently ambiguous expressing the two senses of movement and change of location without invoking the perspective shift. Locative alternations verbs are argued to display different patterns of argument drop, depending on whether they have a lexical specification for the two morphological frames or come to possess one of the two morphological frames derivationally via the perspective shift.

1. Introduction

A number of researchers (Levin and Rapoport 1982, Rappaport and Levin 1985, Pinker 1989 and others) claim that verbs participating in the locative alternation express both motion and change-of-state meanings. With regard to the question of why the locative alternation is invoked by a certain class of verbs, but not others, Pinker (1989) suggests that some verbs are allowed to have two different frames (i.e. motion and change-of-state frames) through the process of 'gestalt shift' (or 'perspective shift'). According to Pinker, the perspective shift can be instantiated if a single event can be viewed in a different way under certain conditions, and this makes it possible for a verb to convey an extra meaning closely related to the one it originally carries.[2] Pinker's main claim on the motivation for the locative alternation is that the shift of a perspective makes it possible for a verb to obtain a derivationally-created frame for arguments, which would not be available otherwise; for instance, if a given verb originally has a frame specifying a

---

[1] This is a revised version of the paper presented at the 7th Mediterranean Morphology Meeting (Nicosia, Cyprus, September, 2009). I am grateful for Geert Booij, Franz Plank and Yoshie Yamamori for comments. I am also thankful to Yu Yile for providing me with Chinese data. The author is responsible for any remaining errors and inadequacies.

[2] Broadly speaking, if the location is conceived of as being 'affected', the locative alternation is made available (see Pinker 1989, Jackendoff 1990 and others). How this type of meaning can be conceived might be subject to language variation. As we will discuss below, in Japanese, this can be achieved, for example, by way of providing a certain meaning invoking 'filling' and 'mountain-like' configurations.

motion event, it can derivationally obtain an additional argument frame for a change-of-state event via the perspective shift.

In Japanese, like many other languages, a certain set of verbs may participate in the locative alternation. We will argue that the perspective shift which Pinker (1989) hypothesizes is empirically justified by looking at Japanese compound verbs. Japanese abounds with compound verbs, and as we will discuss at length below, the grammatical process of verb compounding sometimes induces a change in the possibility of the locative alternation. On the basis of the morphological patterns of arguments obtained by complex predicates, it is shown that in some cases, the locative alternation is indeed created by virtue of the perspective shift, as Pinker claims, but at the same time, it is suggested that in other cases, the locative alternation is not made available by the perspective shift, contrary to Pinker's assumption that one of the two frames of locative alternation verbs is always derived.

Specifically, this paper argues that locative alternation verbs should be divided into the following two classes: (1) the verbs which have both the change-of-state and motion meanings (as their basic meanings), and (2) the verbs which have one of the two meanings as basic, while the other meaning is derived by the perspective shift. The two classes of locative alternation verbs are shown to have distinct syntactic behavior. In Japanese, the arguments of intrinsically ambiguous verbs can be freely dropped in both motion and change-of-state variants as long as their references are inferable from context. In contrast, the omission of the oblique arguments appearing in the variant derived via the perspective shift results in unacceptability. Since accusative arguments are allowed to be dropped in both variants regardless of whether the locative alternation is derived via the perspective shift or not, the Japanese facts make it clear that what is taken to be a sole complement constraint by Pinker (1987) should, in fact, be viewed as a condition on the realization of an oblique argument, which regulates the identification of the morphological pattern of a derived variant.

2. V-V compounding and argument drop

In this section, V-V compounding in Japanese is shown to provide us with insight into the nature of the locative alternation. The data on Japanese compound verbs—formed by combining two verbs—provide one piece of empirical evidence for the claim that some locative alternation verbs acquire the ability to participate in the locative alternation via the perspective shift in the sense of Pinker (1989). In the following discussion, we will show that with the help of the second verb *tumeru* 'fill', the compound verb *siki-tumeru* 'set-fill' is endowed with the ability to participate in the locative alternation, which is not possible with the base verb *siku* 'set'. A close inspection of the Japanese data regarding the compound verb *siki-tumeru* illustrates that locative alternation verbs sometimes come to possess an extra morphological frame derivationally.

To begin, let us discuss how the arguments of locative alternation verbs in Japanese are marked morphologically. First, observe the alternation pattern obtained for a typical locative alternation verb like *nuru* 'smear, paint'.

(1) a. John-ga penki-o kabe-ni nut-ta.

 John-NOM paint-ACC wall-on paint-PAST

 'John smeared paint on the wall.' (Motion)

 b. John-ga kabe-o penki-de nut-ta.

 John-NOM wall-ACC tile-with paint-PAST

 'John smeared the wall with paint.' (Change of State)

Example (1a) shows the morphological pattern of arguments in the 'motion' frame: in (1a), the locative (goal) argument is marked with *ni* 'on, to', and the theme (material) argument—which refers to an entity that moves—is realized as a direct object, marked with accusative case. On the other hand, (1b) is a case involving the 'change-of-state' frame, where the theme argument receives the oblique marking *de* 'with', and the locative argument—which undergoes a change of state—is marked with accusative case. In both variants, the argument referring to an entity that is taken to move or undergo a change is realized as an accusative argument (see Tenny 1994). We assume here, as often claimed (Levin and Rapoport 1982, Rappaport and Levin 1985, Pinker 1989, and others), that verbs participating in the locative alternation possess two different morphological frames, on the grounds that they express two distinct meanings—in this case, one which specifies the movement of some material ('paint') onto a location ('wall') and the other which specifies a change of state which takes place on the location.[3]

Now, for the purpose of confirming the existence of the mechanism of creating the locative alternation, let us discuss how the compound verb *siki-tumeru* 'set-fill'—formed from the verb *siku* 'set' by way of compounding with the second verb *tumeru* 'fill'—behave syntactically. The examples in (2) illustrate that *siki-tumeru* is allowed to participate in the locative alternation, and can take both motion and change-state-of state frames.

(2) a. John-ga tairu-o yuka-ni siki-tume-ta.

 John-NOM tile-ACC floor-on set-fill-PAST

 'John laid the tiles on the floor.' (Motion)

 b. John-ga yuka-o tairu-de siki-tume-ta.

 John-NOM floor-ACC tile-with set-fill-PAST

 'John laid the floor with tiles.' (Change of State)

---

[3] It is important to keep in mind that a locative alternation verb encodes the meaning of change of state pertaining to a location, but not just any type of change. Accordingly, if a locative change is not conceived, no alternation is induced (see Fukui, Miyugawa & Tenny 1985).

The compound verb *siki-tumeru* includes the two verbs *siku* 'set' and *tumeru* 'fill' as its components, so it is necessary to first identify which verb is held responsible for argument realization (see Kageyama 1993). Here, we can state that the compound verb *siki-tumeru* has the first verb *siku* 'set' as its head—i.e. the base verb of the compound verb—because the selectional properties of the arguments of the whole compound verb are determined by the first verb, but not by the second, as can be seen in (3).

(3) a. John-ga   tairu-o   yuka-ni   sii-ta.

      John-NOM   tile-ACC   floor-on   set-PAST

      'John set the tiles on the floor.'

   b. #John-ga tairu-o   yuka-ni   tume-ta.

      John-NOM   tile-ACC   floor-on   fill-PAST

      'John filled the tiles on the floor.'

The verb *siku* expresses a two-dimensional event, and *tumeru* a three-dimensional one (when they are used as independent verbs). Note that what is described by the complex verb *siki-tumeru* in (2) is a sub-type of tile-setting event. Since the setting of tiles on the floor is a two-dimensional event, (3b) is anomalous, where the verb *tumeru* 'fill' is used. In light of this fact, it is reasonable to state that the head of the compound verb *siki-tumeru*, which serves to determine how arguments are realized, should be the first verb *siku*.

The compound verb *siki-tumeru* falls into the 'complementation' type of lexical compound in Kageyama's (1993) analysis. According to Kageyama, lexical compound verbs are classified into the following three types: 'coordinate', 'right-hand head', and 'complementation' types.[4] In Japanese, a compound verb like *naki-sakebu* 'cry-shout' represents a coordinate compound, for the verbs carrying similar meanings are compounded. A compound verb like *hiki-ageru* 'pull-raise' is construed as a right-hand one, in the sense that the first verb acts like an auxiliary verb, specifying a manner of action described by the second verb, whereas its argument structure is determined by the second verb. A compound verb like *arai-ageru* (wash-raise) 'wash up' is classified as the complement type, since the left-hand verb determines the properties of arguments taken by the whole, and the right-hand verb specifies an aspectual meaning. In the case of *siki-tumeru*, the right-hand verb serves to add an aspectual meaning to the left-hand verb, and the left-hand verb determines the argument structure of the whole, hence the complex verb is classified into a 'complementation' type.

Importantly, the base verb *siku* 'set' can take only the theme (i.e. material) as its accusative argument, for the verb describes the movement of the material. Thus, (4a) is acceptable, but (4b) is not.

(4) a. John-ga   tairu-o   yuka-ni   sii-ta.

---

[4] Japanese has syntactic compound verbs, which have analyzable syntactic structures (see Kageyama 1993). In this paper, we do not look at the syntactic compound verbs, since syntactic compounding does not change the morphological patterns of arguments.

> John-NOM   tile-ACC   floor-on   set-PAST
>
> 'John set the tiles on the floor'   (Motion)
>
> b.   *John-ga   yuka-o   tairu-de   sii-ta.
>
> John-NOM   floor-ACC   tile-with   set-PAST
>
> 'John set the floor with tiles.'   (*Change of State)

The compound verb *siki-tumeru* can take the change-of-state frame, which is not available for the simple verb *siku*. A comparison of the data (2) and (4) shows then that when the verb *siku* 'set' is compounded with the verb *tumeru* 'fill', the change-of-state frame is made available, alongside the motion frame, and hence, the compound verb *siki-tumeru* is construed as a locative alternation verb.

Now, the question to be raised here is how the compound verb *siki-tumeru* is furnished with the change-of-state frame, for which the base verb does not have a lexical specification. In the following discussion, we argue that the change-of-state frame is provided derivationally via the perspective shift in the sense of Pinker (1989). The key to understanding this issue lies in a difference in meaning expressed by the base verb and the compound verb. As discussed by a number of researchers (see Rappaport and Levin 1985, Levin and Rappaport Hovav 1995, Pinker 1989, Jackendoff 1990, Kageyama 1980, Okutu 1981, Kishimoto 2001, among many others), verbs participating in the locative alternation (like *paint*) can be assumed to carry two distinct meanings, one which expresses the meaning of moving material onto a location and the other indicates a change of state (pertaining to the location) effected by the movement of the material.

Locative alternation verbs can take the change-of-state frame when the location is conceived of as being affected, most typically, by carrying the sense of 'total' affectedness (Jackendoff 1990, Tenny 1994, Levin and Rappaport 1995). But verbs do not have to carry this 'affected location' meaning when they take the motion frame. The base verb *siku* 'set' does not denote that the whole area of the floor is covered, lacking the 'affected location' sense, so that this verb can only take the motion frame. But once the verb is compounded with *tumeru* 'fill', the meaning that the whole area is covered with tiles is furnished. Thus, we can observe a difference in meaning between the base verb and the compound verb, as illustrated in Figure 1.

siku   *siki-tumeru*

Figure 1: *siku* 'set' and *siki-tumeru* 'set-fill'

The verb *siku* 'set' simply denotes the act of moving some material onto a location, but the compound verb *siki-tumeru*—which has the ability to invoke the locative alternation—additionally carries the 'affected' meaning that the whole area is covered. Since only the motion frame can be provided by the base verb *siku*, we can readily see

that the change-of-state frame (i.e. the *de*-accusative pattern) for *siki-tumeru* is made available via V-V compounding.[5]

Now, bearing the semantic facts of *siku* and *siki-tumeru* in mind, let us now proceed to consider the possibility of argument drop. The verb *siku* 'set' can appear only in the motion frame. As shown in (5), the oblique as well as the accusative argument of *siki-tumeru* can be dropped (provided their reference is recoverable from context).

(5) a. Taro-ga   tairu-o   (yuka-ni) sii-ta.

   Taro-NOM tile-ACC   floor-on   set-PAST

   'Taro set tiles on the floor.'

   b. Taro-ga   (tairu-o) yuka-ni   sii-ta.

   Taro-NOM tile-ACC   floor-on   set-PAST

   'Taro set tiles on the floor.'

By contrast, the compound verb *siki-tumeru* can appear in the change-of-state frame, alongside the motion frame. These two variants show a difference in the possibility of argument drop. In the motion variant, both oblique and accusative arguments can be dropped, as seen in (6).

(6) a. John-ga   tairu-o   (yuka-ni) siki-tume-ta.

   John-NOM tile-ACC   floor-on   set-fill-PAST

   'John laid the tiles on the floor.'   (Motion)

   b. John-ga   (tairu-o) yuka-ni   siki-tume-ta.

   John-NOM tile-ACC   floor-on   set-fill-PAST

   'John laid the tiles on the floor.'   (Motion)

On the other hand, in the change-of-state variant (which constitutes a derived morphological frame that is not available for the base verb), the accusative argument can be omitted, but the omission of the oblique argument results in unacceptability, as seen in (7).

(7) a. Taro-ga   yuka-o   *(tairu-de)   siki-tume-ta.

   Taro-NOM   floor-ACC   tile-with   set-fill-PAST

   'Taro laid the floor with tiles.'   (Change of State)

---

[5] (3a), which contains the verb *siku* 'set', comes to carry a meaning close to the one conveyed by (2a), which includes *siki-tumeru*, if the accusative argument *yuka* 'floor' is replaced with *yuka-itimen* 'all over the floor'. Nevertheless, in this case, the perspective change is not instantiated, since the change of the argument does not affect the meaning of the verb, hence its argument structure. Accordingly, the morphological frame of the verb remains intact, as seen in (i).

(i) a. John-ga   tairu-o   yuka-itimen-ni   sii-ta.
   John-NOM   tile-ACC   floor-all.over-on   set-PAST
   'John set the tiles on the floor'.   (Motion)

   b. *John-ga   yuka-itimen-o   tairu-de   sii-ta.
   John-NOM   floor-all.over-ACC   tile-with   set-PAST
   'John set the floor with tiles.'   (*Change of State)

b. John-ga   (yuka-o)   tairu-de   siki-tume-ta.

John-NOM floor-ACC tile-with   set-fill-PAST

'John laid the floor with tiles.'   (Change of State)

Here, an asymmetry emerges in the possibility of argument omission: in the motion variant, the two internal arguments can be dropped without any problem; in the change-of-state variant, the omission of the oblique argument results in unacceptability, but no problem arises when the accusative argument is dropped.

In discussing why certain arguments selected by locative alternation verbs are prevented from undergoing deletion, it is instrumental to refer to Pinker's (1989) analysis on the locative alternation. According to Pinker (1989), locative alternation verbs invoke reference to two different semantic structures. The locative alternation emerges, for instance, when a verb can have the two semantic structures "x causes y to move to z" and "x causes z to change its state by means of moving y to z." In Pinker's theory, these meanings are related by a lexical rule, and this relation is established by what he calls 'gestalt shift' (i.e. the perspective shift). The perspective shift is a way of viewing the same event differently—a reinterpretation of an event from a different perspective, which can be instantiated in some context. Thank to this perspective shift, certain verbs are allowed to acquire related meanings derivationally, and hence additional argument frames, which would otherwise not be available for the verbs, become available.

Pinker claims that whether or not the perspective change is instantiated on locative alternation verbs can be determined by looking at what can stand as a sole complement, on the basis of examples like (8).

(8) a.   He piled the books (on the shelf).

   b.   He piled the shelf *(with the books).

Pinker's assumption is that the variant which supplies the obligatory argument of the verb as its object is basic. In the case of *pile*, the omission of the PP is not allowed in the change-of-state frame, so the verb should be lexically specified for the motion frame (8a), the change-of-state frame (8b) being derived via the perspective shift.[6]

Note, however, that by just looking at English examples, it is not so easy to ascertain whether or not Pinker's directionality claim for the locative alternation is appropriate. This is precisely because, in English, arguments are often not omissible even if the perspective shift is not invoked. Even though the omission of internal arguments selected by three-place verbs is sometimes allowed (see Pesetsky 1995, Dowty 1979, and others), there are nevertheless cases in which their omission results in unacceptability, as (9) illustrates.

(9) a.   John put the book *(on the table).

   b.   John put *(the book) on the table.

---

[6] The facts of argument drop have often been reported in the literature (e.g. Fraser 1971, Rappaport and Levin 1985, Levin 1986), but Pinker (1989) is the first to advance a theoretical claim on the constraint of argument drop.

Needless to say, the verb *put* does not invoke any alternation, but still, the omission of the locative PP *on the table* results in unacceptability. Even though the PP must be expressed, the verb *put* does not have a variant where the locative argument appears as an object. This suggests that the presence of a selectional restriction imposed on English verbs often precludes us from testing for Pinker's hypothesis.[7] In Japanese, by contrast, any argument may, in principle, be missing insofar as its reference is recoverable from context, as exemplified in (10).

(10) a. Taro-ga   (yuka-ni)   tairu-o   oi-ta.

      Taro-NOM   floor-on   tile-ACC   put-PAST

      'Taro put tiles on the floor.'

    b. Taro-ga   yuka-ni   (tairu-o)   oi-ta.

      Taro-NOM   floor-on   tile-ACC   put-PAST

      'Taro put tiles on the floor.'

The verb *oku* 'put' in (10), just like the English verb *put* in (9), takes an accusative, as well as an oblique argument, but unlike English, the two arguments selected by *oku* do not have to be overtly realized (on the condition that their reference may be fixed contextually). Since Japanese is free from the restriction concerning the overt realization of arguments that is often found in English, we can say that Japanese offers the setting suitable to check the nature of argument drop associated with the locative alternation.

To return, recall that the locative alternation verb *siki-tumeru* (set-fill) does not allow the omission of the oblique argument in the change-of-state frame variant, as in (7a). Moreover, its base verb can only take the motion frame, but the compound verb allows for the locative alternation, on the grounds that the compound verb comes to express the sense that the location is completely affected. The compound verb examples show that the derived frame, which the base verb does not possess, is provided when the predicate expresses the sense of 'total affectedness' pertaining to the location, as seen above. From these facts, we can confirm that Pinker's proposal on the correlation between the directionality of derivation and the possibility of argument omission is essentially on the right track.[8] Obviously, the directionality of a frame derivation for the compound verb *siki-tumeru* is from the motion variant to the change-of-state one.

The data regarding the omission of arguments with the Japanese compound *siki-tumeru* also show that while the oblique argument of a derived frame cannot be dropped, the acceptability of the derived frame is not affected by the omission of accusative

---

[7] Goldberg (1995: 177) points out some English examples where neither the theme nor the location stands as a sole complement.

 (i) a.   Pat heaped mashed potatoes * (onto her plate).

    b.   Pat heaped her plate *(with mashed potatoes).

In this case, the strategy of argument drop does not provide a measure for the directionality of the derivation of morphological frames (see also Iwata 2008). We assume that this arises from a selectional restriction imposed on the verb, rather than a constraint on the perspective shift.

[8] In Japanese, the derivation of the locative alternation goes from 'motion' to 'change of state', and as far as I can see, there is no clear case displaying the opposite directionality of derivation.

arguments. The Japanese fact suggests then that the directionality of derivation is not measured by just checking what argument is obligatorily expressed, contrary to Pinker's hypothesis. This raises a question why the oblique argument (i.e. the theme) must be expressed in the derived change-of-state frame. The fact follows straightforwardly, given that only an oblique argument can be used as a syntactic clue to identify the morphological frame of the clause.[9]

To make matters concrete, consider the case in which the compound verb *siki-tumeru* appears in the derived change-of-state frame, as in (7a). In (7a), if the oblique argument (i.e. *tairu-de* ['tile-with']) is provided, we can visibly identify the argument alignment for the derived frame. However, if this oblique argument is not realized, it is not possible to identify the use of the derived frame uniquely in syntactic terms (i.e. we cannot tell whether or not the argument alignment is based on the derived frame). This suggests that in order for the verb to obtain a derived frame, a syntactic condition is imposed, such that the use of a derived frame is uniquely identified syntactically by overtly expressing the oblique argument. In the non-derived frame, on the other hand, no such peculiar restriction is imposed, for both oblique and accusative arguments of *siki-tumeru* appearing in the motion frame can be omitted without causing any problem, as seen in (6). If the frame for which the verb has an intrinsic lexical specification is used when an oblique argument is absent, it falls out that the omission of the oblique argument in the derived frame results in unacceptability.

The proposal advanced here amounts to saying that two conditions need to be satisfied in order for the perspective shift to be successful. One condition is semantic, requiring that the meaning of an affected location be supplied in some way; in the case of *siki-tumeru*, this semantic condition is fulfilled by virtue of compounding with predicate *tumeru*, since the second verb allows the compound verb to obtain the additional meaning that the entire area is covered. The other condition concerns the syntactic identification of the derived frame, which can be fulfilled by way of overtly expressing the oblique argument marked with *de* 'with'.

(11)   [PP   OBJ   V1-V2]

the meaning of 'affected location' supplied

the identification of a syntactic frame

The data regarding the compound verb *siki-tumeru* illustrate that the semantics alone is not sufficient for a perspective change to take effect, but that the syntactic identification of the derived morphological frame is also necessary.

Incidentally, the second verb *tumeru* included in the compound verb allows for the locative alternation when used as an independent verb.

(12) a. Taro-ga   hon-o   hondana-ni   tume-ta.

---

[9] Since English does not offer the data that lead to this generalization, it is not too surprising that Pinker does not provide any specific proposal on this phenomenon.

Taro-NOM book-ACC bookcase-on stuff-PAST

'Taro stuffed books into the bookcase.' (Motion)

b. Taro-ga hon-de hondana-o tume-ta.

Taro-NOM book-with bookcase-ACC stuff-PAST

'Taro stuffed the bookcase with books.' (Change of State)

Given this fact, one might be tempted to say that the second verb supplies the compound verb *siki-tumeru* with the change-of-state frame. This is not the case, however, as can be ascertained if we look at the possibility of argument omission. The examples show that *siki-tumeru* allows the oblique argument to be omitted in both change-of-state and motion variants without affecting their acceptability.

(13) a. Taro-ga hon-o (hondana-ni) tume-ta.

Taro-NOM book-ACC bookcase-on stuff-PAST

'Taro stuffed books into the bookcase.' (Motion)

b. Taro-ga (hon-de) hondana-o tume-ta.

Taro-NOM book-with bookcase-ACC stuff-PAST

'Taro stuffed the bookcase with books.' (Change of State)

Needless to say, for this predicate, the accusative argument as well can be dropped in both change-of-state and motion variants, as shown in (14).

(14) a. Taro-ga (hon-o) hondana-ni tume-ta.

Taro-NOM book-ACC bookcase-on stuff-PAST

'Taro stuffed books into the bookcase.' (Motion)

b. Taro-ga hon-de (hondana-o) tume-ta.

Taro-NOM book-with bookcase-ACC stuff-PAST

'Taro stuffed the bookcase with books.' (Change of State)

This pattern of argument drop differs from what we observe for the compound verb *siki-tumeru*. The prohibition against dropping the oblique argument should be obtained when the morphological frame which the verb does not inherently possess is provided by the perspective shift. If so, it is reasonable to state that the second verb *tumeru* does *not* provide the change-of-state frame, which is available for *siki-tumeru* (see also section 3).

The peculiar alternation pattern of compound verbs invoking the locative alternation is observed not only in Japanese but also in Chinese. In Chinese, just like Japanese, V-V compounds are formed quite productively. Furthermore, Chinese patterns with Japanese, in that the language allows for the omission of arguments freely if their reference can be inferred from context (Huang and Li 1996). Owing to these properties, we can easily present another case from Chinese which illustrates the creation of the locative alternation via the perspective shift.

To be concrete, let us consider how the verb *sē* 'stuff' and the V-V compound verb *sē mǎn* 'stuff-full' behave syntactically. In the first place, the verb *sē* 'stuff' can take a motion frame, but not a change-of-state one when it is used in isolation, as seen in (15).

(15) a. *Zhāngsān yòng   shū   sē   le   bāo.
Zhangsan   with   book   stuff ASP   bag
'Zhangan stuffed the bag with books.'   (Change of State)

    b. Zhāngsān   zài bāo lǐ sē   le shū.
Zhangsan   at   bag in stuff   ASP book
'Zhangsan stuffed books into the bag.'   (Motion)

The examples suggest that the Chinese verb *sē* 'stuff' can only be used as a motion verb. In the second, Chinese makes a change-of-state variant available if the verb *sē* 'stuff' is compounded with another verb *mǎn* 'full'. The following examples illustrate that both the change-of-state and the motion frames are available with the complex verb *sē mǎn* 'stuff-full'.

(16) a. Zhāngsān yòng shū   sē   mǎn le   bāo.
Zhangsan with   book   stuff full ASP   bag
'Zhangsan stuffed the bag with books.'   (Change of State)

    b. Zhāngsān gěi   bāo   sē   mǎn le   shū.
Zhangsan to   bag   stuff   full ASP   book
'Zhangsan stuffed books into the bag.'   (Motion)

It goes without saying that there is a difference in meaning between the simple verb *sē* 'stuff' and the complex predicate *sē mǎn* 'stuff-full', since the latter, but not the former, indicates that the bag (i.e. location) is full. This suggests that the meaning of affectedness on the location to invoke the locative alternation is not obtained for the simple verb *sē*. If so, it is reasonable to state that in Chinese, the change-of-state variant for *sē mǎn* (16b) is made available via V-V compounding, which allows us to view the location as being totally affected.

The Chinese compound verb *sē mǎn* 'stuff-full' provides a clear case in support of the present view that an extra argument frame can be provided via the perspective shift, because neither of the component verbs has a lexical specification for the change-of-state variant.

(17) a. Bāo/*Shū mǎn   le.
bag/book   full   ASP
'The bag is full/the books are full.'

  b. *Bāo   yòng shū   mǎn le.
bag   with   book   full ASP
'The bag is full with books.'

As shown in (17a), the predicate *mǎn* 'full' can take a location subject when it appears in isolation. Nevertheless, this predicate cannot appear in the change-of-state frame, as

seen in (17b). This shows that no change-of-state variant is available for any components of the complex predicate *sē mǎn* 'stuff-full'.

If the locative alternation is created by way of the perspective shift, we would naturally expect that the change-of-state variant—which should count as a derived frame—does not allow the obliquely-marked argument to be omitted. This is in fact the case. The examples in (18) illustrate that in both motion and change-of-state variants formed from *sē mǎn*, the direct objects may be dropped without affecting their acceptability.

(18) a. Zhāngsān yòng shū sē mǎn le (bāo).
        Zhangsan with book stuff full ASP bag
        'Zhangsan stuffed the bag with books.'   (Change of State)

    b. Zhāngsān gěi bāo sē mǎn le (shū).
        Zhangsan to bag stuff full ASP book
        'Zhangsan stuffed books into the bag.'   (Motion)

If we look at the omission of the oblique arguments, we find a contrast in acceptability between the two variants. As shown in (19), the change-of-state variant does not allow the oblique argument to be dropped, while the motion variant does.

(19) a. Zhāngsān */??(yòng shū) sē mǎn le bāo.
        Zhangsan with book stuff full ASP bag
        'Zhangsan stuffed the bag with books.'   (Change of State)

    b. Zhāngsān (gěi bāo) sē mǎn le shū.
        Zhangsan to bag stuff full ASP book
        'Zhangsan stuffed books into the bag.'   (Motion)

The patterns observed above with regard to the omission of the oblique arguments are in conformity with the claim advanced about the perspective shift. In Chinese as well as in Japanese, certain compound verbs are endowed with the ability to induce the locative alternation via the perspective shift. If the locative alternation is created via the perspective shift, the obliquely-marked argument in the derived frame cannot be dropped.

In this section, by looking at some cases where the locative alternation is created by V-V compounding, we have suggested that the oblique argument appearing in the frame derived by the perspective shift cannot be dropped, owing to the condition on its syntactic identification. In the next section, we will turn to different types of complex predicates, and show that their data lend further empirical support to the proposed analysis taking some locative alternation verbs to be derived via the perspective shift.

3. Another type of complex predicate inducing locative alternation

With regard to the omission of arguments, the same pattern of distribution that we observe for the complex verb *siki-tumeru* is found in a complex verb like *yama-mori-ni suru* 'make a mountain-like serving'. The complex predicate *yama-mori-ni suru* is

formed by adding *suru* 'make' to the complex nominal expression *yama-mori*, where a noun *yama* 'mountain' is compounded with the nominalized *mori* 'serving'. As seen in (20), this complex predicate can invoke the locative alternation.

(20) a. John-ga   otyawan-ni   gohan-o   yama-mori-ni si-ta.

      John-NOM bowl-on   rice-ACC   mountain-serve make-PAST

      (Lit.) 'John served rice on the bowl like a mountain.'   (Motion)

    b. John-ga   otyawan-o   gohan-de   yama-mori-ni   si-ta.

      John-NOM   bowl-ACC   rice-with   mountain-serve make-PAST

      (Lit.) 'John served the bowl with rice like a mountain.'   (Change of State)

Here, we observe the following pattern with regard to the omission of arguments.

(21) a. John-ga   (otyawan-ni) gohan-o   yama-mori-ni   si-ta.

      John-NOM bowl-on   rice-ACC   mountain-serve make-PAST

      (Lit.) 'John served rice on the bowl like a mountain.'   (Motion)

    b. John-ga   otyawan-o *(gohan-de)   yama-mori-ni   si-ta.

      John-NOM   bowl-ACC   rice-with   mountain-serve   make-PAST

      (Lit.) 'John served the bowl with rice like a mountain.'   (Change of State)

Since the oblique argument cannot be elided in the change-of-state variant, we can assume that this variant is derived by the perspective shift. As we will discuss below, in (20), the presence of a noun like *yama* 'mountain' makes it possible for the complex predicate to invoke the locative alternation.

Note that the head of the complex predicate—which provides the motion frame of arguments—should be the deverbal noun *mori* 'serving'. This state of affairs is naturally expected, since the event described by the predicate *yama-mori-ni suru* is regarded as one sort of 'serving' event, which the verb *moru* 'serve' should denote. (22) shows that the verb *moru* can take only the motion frame.

(22) a. John-ga   otyawan-ni   gohan-o   mot-ta.

      John-NOM   bowl-on   rice-ACC   serve-PAST

      'John served rice on the bowl.'   (Motion)

    b. *John-ga   otyawan-o   gohan-de mot-ta.

      John-NOM   bowl-ACC   rice-with serve-PAST

      'John served the bowl with rice.'   (*Change of State)

The fact that the verb *moru* has a lexical specification only for the motion frame can further be confirmed by (23).

(23) a. John-ga   tyawan-ni gohan-o   oo-mori-ni   si-ta.

      John-NOM bowl-on   rice-ACC   large-serve   make-PAST

      (Lit.) 'John served rice on the bowl.'   (Motion)

    b. *John-ga   tyawan-o gohan-de   oo-mori-ni si-ta.

      John-NOM   bowl-ACC rice-with   large-serve make-PAST

(Lit.) 'John served the bowl with rice.'   (*Change of State)

As seen in (23), if the noun *yama-* 'mountain' is replaced with another noun *oo-* 'large', the locative alternation is no longer available. The complex predicate *oo-mori-ni suru* displays exactly the same morphological pattern as observed by the verb *moru* 'serve', suggesting that the base verb is lexically specified only for the motion frame.[10]

At first blush, the complex predicates *yamo-mori-ni suru* and *oo-mori-ni suru* look like expressing similar meanings, but if we look at the difference in meaning between the two, we can ascertain that these verbs fall into distinct classes. As illustrated in Figure 2, *oo-mori* 'large serving' specifies just the quantity of served food, but *yama-mori* 'mountain-like serving' indicates the shape of the food (served on a bowl) regardless of its quantity.

oo-mori 'large' serving   *yama-mori* 'mountain' serving
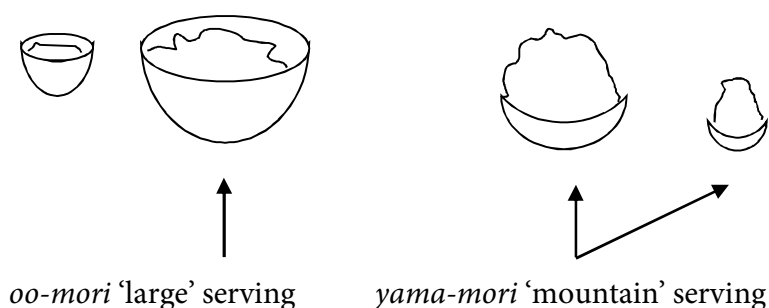
Figure 2: *oo-mori* and *yama-mori*

In this particular case, the noun *yama*—which invokes a conceptualization of mountain-like shape—provides the meaning of totally affected location. The component of meaning indicating a mountain-like shape helps instantiate the perspective shift in the sense of Pinker. Consequently, only the verb *yama-mori-ni suru*, but not *oo-mori-ni suru*, participates in the locative alternation. Since the event described by the verb *moru* 'serve' represents a motional one, the change-of-state frame should be provided by way of the perspective shift. Accordingly, we can naturally anticipate that *yama-mori-ni suru* should not allow for the omission of the oblique argument appearing in the change-of-state frame.

We can readily confirm that the noun *yama* 'mountain' included in the complex predicate *yama-mori-ni suru* 'make a mountain-like serving' does not provide the

---

[10] Both arguments are omissible with the complex predicate *oo-mori-ni suru* 'make a large serving' as illustrated in (i).

(i) a.   John-ga   otyawan-ni   (gohan-o)   oo-mori-ni   si-ta.
        John-NOM   bowl-on   rice-ACC   large-serve   make-PAST
        (Lit.) 'John served rice on the bowl.'

    b.   John-ga   (otyawan-ni)   gohan-o   oo-mori-ni   si-ta.
        John-NOM   bowl-on   rice-ACC   large-serve   make-PAST
        (Lit.) 'John served rice on the bowl.'

The predicate *oo-mori-ni suru* does not involve locative alternation. Since the verb has a lexical specification for the frame in (i), the omission of the arguments is allowed without problem.

change-of-state frame, since *yama* cannot be used as a predicate describing a change of state pertaining to a location, as exemplified in (24).

(24) *John-ga   otyawan-o gohan-de yama-ni   si-ta.

John-NOM   bowl-ACC   rice-with mountain-DAT make-PAST

'John made the rice a mountain with a bowl.'   (*Change of State)

Note also that the verb *suru* 'make' does not play a role for determining the two morphological frames. This can be seen by the fact that the intransitive versions of *yama-mori-ni suru* and *oo-mori-ni suru* contain the verb *naru* 'become' rather than *suru* 'make', as in *yama-mori-ni naru* and *oo-mori-ni naru*[11]. The fact indicates that no components of the complex predicate can provide the derivational change-of-state frame specifying an affected location. As we have seen above, since the deverbal noun *mori* (< *moru* 'serve') in the complex predicate specifies only the motion frame, the observed pattern of argument drop should be derived from the fact that the change-of-state frame is created by the perspective shift.[12]

It is also worth noting here that the perspective shift can be effected even for a verb without compounding in Japanese. This is illustrated in (25).

(25) a. John-ga   (ude-ni) hootai-o   mai-ta.

John-NOM arm-on   bandage-ACC   roll-PAST

'John rolled a bandage around his arm.'   (Motion)

b. John-ga   *(hootai-de)   ude-o   mai-ta.

John-NOM bandage-with   arm-ACC   roll-PAST

'John rolled the arm with a bandage.'   (Change of State)

The data in (25) suggest that compounding is not necessarily required to instantiate the perspective shift, insofar as a single verb may satisfy a certain semantic restriction to effect this change. I will turn to the discussion of this point in the next section.

---

[11] The intransitive version of the complex predicate is created if the verb *naru* 'become' (or copula *da*) is added instead of the verb *suru* 'make'.

(i) a.   Otyawan-ni gohan-ga   yama-mori-ni   nat-ta.

bowl-on   rice-NOM   mountain-serve become-PAST

'Rice was served on the bowl like a mountain.'   (Motion)

b.   Otyawan-ga gohan-de   yama-mori-ni   nat-ta.

bowl-NOM   rice-with   mountain-serve become-PAST

'The bowl was served with rice like a mountain.'   (Change of State)

Needless to say, the intransitive complex predicate behaves in exactly the same way as its transitive counterpart with regard to the alternation of argument frames and argument drop.

[12] Other complex verbs which pattern with the alternating *yamamor-ni suru* versus the non-alternating *oo-mori-ni suru* include the pair of *yama-zumi-ni suru* (mountain-pile-DAT make) 'pile up' *hira-zumi-ni suru* (flat-pile-DAT make) 'display'. The verb *yama-sumi-ni suru* can instantiate the locative alternation, but *hira-zumi-ni suru* does not. Here again, a mountain-like configuration is a crucial factor to determine the possibility of the locative alternation.

4. Compounding with ambiguous verbs

The present analysis predicts that the locative alternation verbs which are inherently ambiguous should impose no restriction on the omission of arguments in both variants. In fact, in Japanese, many locative alternation verbs should be ambiguous according to this criterion; i.e. many locative alternation verbs allow their arguments to be dropped freely. For instance, as shown in (26), with the verb *nuru* 'paint', the oblique argument can be omitted in both change-of-state and motion variants, provided their reference is recoverable from the context.

(26) a. Taro-ga   (kabe-ni)   akai   penki-o   nut-ta.

Taro-NOM   wall-on   red   paint-ACC   paint-PAST

'Taro painted red paint on the wall.'   (Motion)

b. Taro-ga (akai   penki-de)   kabe-o   nut-ta.

Taro-NOM red   paint-with   wall-ACC   paint-PAST

'Taro painted the wall with red paint.'   (Change of State)

In a similar vein, the accusative-marked argument may be dropped in the two variants, as shown in (27).

(27) a. Taro-ga   kabe-ni   (akai penki-o)   nut-ta.

Taro-NOM   wall-on   red   paint-ACC   paint-PAST

'Taro painted red paint on the wall.'   (Motion)

b. Taro-ga   akai penki-de   (kabe-o)   nut-ta.

Taro-NOM   red   paint-with   wall-ACC   paint-PAST

'Taro painted the wall with red paint.'   (Change of State)

In both change-of-state and motion variants, acceptability is not affected even if the oblique argument is omitted. If the derivational direction invoked by the perspective shift can be measured with recourse to argument omission, it is reasonable to say that both motion and change-of-state frames for the verb *muru* 'paint' should not be provided derivationally.

Pinker is aware of the fact that in English some locative alternations allow their oblique arguments (i.e. PPs) to be omitted in both the variants.

(28) a. He loaded the gun (with the bullets).

b. He loaded the bullets (onto the gun).

On the assumption that one of the two frames of the locative alternation verbs is always created by the perspective shift, Pinker argues that in a case like (28), the directionality of derivation can be determined by considering which variant expresses a complete thought if the PP is omitted. Pinker reports that *He loaded the gun* is felt to be a complete thought, whereas *He loaded the bullets* sounds like a truncated sentence. Thus, for Pinker, the motion frame for the verb *load* should be a derived one.

The Japanese fact suggests that Pinker's analysis is not warranted, however. As we have discussed in section 2, the facts of the compound verb *siki-tumeru* show that for the

perspective shift to be successful, the presence of the derived frame is syntactically identified, and thus, the presence or absence of the perspective shift can be measured by looking at whether or not an oblique argument can be omitted. In Japanese, arguments can be dropped freely if the frame in which the verb appears is not created derivationally via the perspective shift. As a matter of fact, the possibility of argument omission does not change whether the sentence is felt to be complete or not. This suggests that there is a qualitative difference between the verbs which tolerate the omission of oblique arguments and the ones which do not.

In the light of this consideration, it is plausible enough to hypothesize that the verb should be inherently ambiguous when oblique arguments can be successfully omitted in both variants. In what follows, we will provide an argument in support of the claim that some locative alternation verbs should be inherently ambiguous—i.e. they are equipped with both motion and change-of-state meanings (as their basic meanings).

To make the point, let us first look at the behavior of compound verbs formed from the verb *nuru* 'paint' (as the base verb). The verb *nuru* can be compounded with a number of different verbs. What is notable about verb compounding with *nuru* is that either of the two variants is made unavailable with a certain choice of the second verb. For instance, the compound verb *nuri-tukeru* 'paint-attach' can only take the motion frame, as seen in (29).

(29) a. Taro-ga   (kabe-ni) akai penki-o   nuri-tuke-ta.
       Taro-NOM wall-on   red paint-ACC paint-attach-PAST
       'Taro painted red paint on the wall.'   (Motion)

    b. *Taro-ga   akai penki-de kabe-o   nuri-tuke-ta.
       Taro-NOM red paint-with wall-ACC   paint-attach-PAST
       'Taro painted the wall with red paint.'   (*Change of State)

The change-of-state variant (29b) is not available for the compound verb *nuri-tukeru*, and as indicated in (29a), the oblique argument in the motion variant can be omitted. On the other hand, the compound verb *nuri-ageru* 'paint-up', where the second verb *ageru* 'up' conveys the meaning of 'completeness', can take only the change-of-state frame, as seen in (30).

(30) a. *Taro-ga   kabe-ni akai penki-o   nuri-age-ta.
       Taro-NOM wall-on red   paint-ACC   paint-up-PAST
       'Taro painted up red paint on the wall.'   (*Motion)

    b. Taro-ga (akai   penki-de) kabe-o   nuri-age-ta.
       Taro-NOM red   paint-with wall-ACC   paint-up-PAST
       'Taro painted up the wall with red paint.'   (Change of State)

With the compound verb *nuri-ageru*, the motion variant is not available, as indicated in (30a), and (30b) shows that the oblique argument in the change-of-state variant can be omitted.

V-V compounding sometimes does not affect the possibility of the locative alternation. The compound verb *nuri-takuru* 'daub', just as in the simple verb *nuru* 'paint', allows for the locative alternation, as illustrated in (31).

(31) a. Taro-ga   (kabe-ni) akai   penki-o   nuri-takut-ta.

Taro-NOM   wall-on   red   paint-ACC paint-daub-PAST

'Taro daubed red paint on the wall.'   (Motion)

 b. Taro-ga   (akai   penki-de) kabe-o   nuri-takut-ta.

Taro-NOM   red   paint-with wall-ACC paint-daub-PAST

'Taro daubed the wall with red paint.'   (Change of State)

The second verb specifies a manner of 'painting', and the possibility of the locative alternation remains unaffected. The important point is that some variants of the locative alternation verbs are rendered unavailable via the process of V-V compounding.

In cases where compounding eliminates either of the two variants, we can postulate that the choice of the morphological frames possessed by the first base verb is restricted by virtue of its semantic incongruity with the second verb. The change-of-state variant indicates a change that takes place on a location, which arises from the transitory process of moving material onto the location. The part of the meaning of the motion variant indicating a transitory process is not compatible with the completive meaning expressed by *ageru*, hence the motion variant is not available for the compound verb *nuri-ageru* (see Kishimoto 2001). On the other hand, when *nuru* 'paint' is compounded with *tukeru*, the second verb adds to the base verb the meaning which indicates the movement of the material. This brings out the consequence of eliminating the change-of-state variant, because the component of meaning which expresses a change of state pertaining to the location becomes incompatible. Given that the compound verbs *nuri-tukeru* and *nuri-ageru* eliminate some of the frames available for the base verb, we can state that the verb *nuru* should have a lexical specification for both motion and change-of-state frames.

Now, for the purpose of providing further justification for our proposal that some locative alternation verbs should be inherently ambiguous, we will take a look at another class of alternation verbs, which include verbs expressing tying and sewing such as *musubu* 'tie' and *sibaru* 'bind'[13]. These verbs can easily give rise to compound verbs, and participate in the locative alternation. Notably, the verbs are divided into two subclasses depending on whether or not they involve the perspective shift. One class includes verbs like *musubu* 'tie', *kukuru* 'tie', and the other *sibaru* 'bind' and *yuu* 'tie' (in addition, *nuu* 'sew', *amu* 'knit', and so forth).

---

[13] Apparently, this class of verbs does not show a partitive/holistic effect observed for other typical locative alternation verbs. Nevertheless, the verbs can be thought of as falling into a sub-class of locative alternation verbs, since they display the same morphological patterns as other locative alternation verbs.

First, the verb *musubu* 'tie' allows for an alternation between the change-of-state and the motion variants.[14]

(32) a. John-ga   hako-ni   himo-o   musun-da.

      John-NOM box-on   string-ACC tie-PAST

      'John tied a string around the box.'   (Motion)

    b. John-ga   himo-de   hako-o   musun-da.

      John-NOM   string-with   box-ACC tie-PAST

       'John tied the box with a string.'   (Change of State)

With this verb, the two variants display a difference with regard to the omission of the oblique arguments, as (33) illustrates.

(33) a. John-ga   (hako-ni) himo-o   musun-da.

      John-NOM box-on   string-ACC tie-PAST

      'John tied a string around the box.'   (Motion)

    b. John-ga   *(himo-de)   hako-o   musun-da.

      John-NOM string-with   box-ACC tie-PAST

      'John tied the box with a string.'   (Change of State)

In the change-of-state variant, the oblique argument is not omissible, but in the motion variant, the oblique argument can be dropped.

The verb *sibaru* 'bind', just like *musubu* 'tie', can take both the motion and change-of-state frames, as in (34).

(34) a. John-ga   hako-ni   himo-o   sibat-ta.

      John-NOM box-on   string-ACC bind-PAST

      'John bound a string around the box.'   (Motion)

    b. John-ga   himo-de   hako-o   sibat-ta.

      John-NOM string-with   box-ACC   bind-PAST

      'John bound the box with a string.'   (Change of State)

Despite the fact that the verbs *sibaru* and *musubu* allow for the locative alternation, *sibaru* displays syntactic behavior that crucially differs from *musubu*. For this verb, the oblique argument can be omitted in both variants, as seen in (35).

(35) a. John-ga   (hako-ni) himo-o   sibat-ta.

      John-NOM box-on   string-ACC bind-PAST

      'John bound a string around the box.'   (Motion)

    b. John-ga   (himo-de)   hako-o   sibat-ta.

---

[14] In the locative alternation examples involving *musubu* 'tie', *sibaru* 'bind' and the compound verbs formed from these verbs, the arguments *himo* 'string' and *hako* 'the box' should be understood to represent the theme and the location, respectively. Other interpretations might be possible in some cases, but they are not relevant for the present purposes.

John-NOM string-with   box-ACC   bind-PAST

'John bound the box with a string.'   (Change of State)

The fact suggests that the change-of-state frame for *musubu* 'tie', but not *sibaru* 'bind', is derived by the perspective shift: that is, in the case of the locative alternation involving *musubu*, the change-of-state variant has a derived frame.

We can postulate that the difference in the possibility of argument drop is derived from a difference in meaning expressed by the two verbs: the verb *musubu* describes a change that takes place on the material; that is, the event described by *musubu* indicates that the string is tightened up, but the verb does not specify whether or not the location is affected (i.e. the verb does not say anything about the tightness of the box). On the other hand, verb *sibaru* describes an event in which both the material and the location are affected (by tightening). For *sibaru*, then, both the material and the location are conceived of as undergoing a change, which suggests that the verb should have a lexical specification for the motion and the change-of-state frames. In the case of *musubu*, since the verb only means that the material is tightened, we can state that the motion frame, but not the change-of-state frame, is lexically specified for the verb, and that the change-of-state frame (32b) should be supplied via the perspective shift. Thus, as seen in (33), an asymmetry arises with regard to argument omission. Since the changes taking place on the material and the location should have a fairly close causal relation, we can assume that the perspective change can be invoked for the verb *musubu* even when it is used independently (without compounding).

Compound verbs can be easily formed in this class of verbs. Let us now look at what syntactic behavior the verbs *sibaru* and *musubu* show with regard to argument omission. First, the verb *sibaru* can be combined with *tukeru*, in which case only the motion variant is made available, as illustrated in (36).

(36) a. John-ga   himo-o   (hako-ni) sibari-tuke-ta.

      John-NOM string-ACC box-on   bind-attach-PAST

      'John bound a string around the box again.'   (Motion)

    b. *John-ga   himo-de   hako-o   sibari-tuke-ta.

      John-NOM   string-with box-ACC   bind-attach-PAST

      'John bound the box with a string again.'   (*Change of State)

The second verb *tukeru* 'attach' adds a motional meaning to the base verb *sibaru*, and this brings out the consequence that the change-of-state frame is eliminated from the compound verb, even though the base verb can take this frame. Accordingly, only the motion variant is available with the compound verb *sibari-tukeru*. In the motion variant, the oblique argument can be omitted without affecting its acceptability.

Second, by way of compounding *sibaru* with the verb *ageru*, which carries a 'completive' meaning, it is possible to form the change-of-state variant, but the other variant—the motion variant—is excluded, as shown in (37).

(37) a. *John-ga   himo-o   hako-ni   sibari-age-ta.

John-NOM   string-ACC   box-on   bind-up-PAST

'John bound a string around the box.'   (*Motion)

b. John-ga   (himo-de)   hako-o   sibari-age-ta.

John-NOM   string-with   box-ACC   bind-up-PAST

'John bound the box with a string.'   (Change of State)

The addition of the verb *ageru* 'up' to the base verb via compounding brings out the consequence that the motional meaning is eliminated from the verb. In the change-of-state variant, the oblique argument can be omitted.

Turning now to the case of *musubu*, the change-of-state variant is eliminated when the verb *tukeru* is compounded with it, as in (38).

(38) a. John-ga   (hako-ni)   himo-o   musubi-tuke-ta.

John-NOM   box-on   string-ACC tie-attach-PAST

'John tied a string around the box.'   (Motion)

b. *John-ga   himo-de   hako-o   musubi-tuke-ta.

John-NOM   string-with box-ACC tie-attach-PAST

'John tied the box with a string again.'   (*Change of State)

The observed pattern of argument realization in (38) is naturally expected, since the second verb *tukeru* compounded to the base verb *musubu* codes a motional meaning. Therefore, the compounding of *musubu* with *tukeru* should have the effect of picking out the motion variant. Interestingly, in the case of the compound verb *musubi-ageru*, where the second *ageru* should carry a 'completive' meaning, both the motion and the change-of-state variants are not acceptable, as seen in (39).

(39) a. *John-ga   hako-ni   himo-o   musubi-age-ta.

John-NOM   box-on   string-ACC   tie-up-PAST

'John tied a string around the box.'   (*Motion)

b. *John-ga   hako-o   himo-de   musubi-age-ta.

John-NOM   box-ACC   string-with   tie-up-PAST

'John tied the box with a string.'   (*Change of State)

The absence of the motion variant would be expected, since the addition of the second verb *ageru* 'up' brings out the consequence that the motion meaning is eliminated from the verb. In the case of the compound verb *musubi-ageru*, the change-of-state variant is absent as well, however, even though this variant should be semantically compatible with the completive *ageru*, as seen in the case of *sibari-ageru* in (37b).

The fact might look puzzling at first, but given that the base verb *musubu* is not lexically specified to take the change-of-state frame, it naturally follows that the variant is not available for *musubi-ageru*. Recall that the primary syntactic effect obtained by compounding with the completive *ageru* is to eliminate the change-of-state frame available for the base verb. Note that *musubu* does not allow the oblique argument in the change-of-state variant to be dropped, which suggests that the motion frame be

provided by invoking the perspective shift. If the base verb *musubu* inherently lacks a lexical specification for the change-of-state frame, there is a sense in which the compound verb *musubi-ageru* cannot appear in the change-of-state frame.

This does not mean that the compound verb *musubi-ageru* does not exist in the lexical entries of the language. In effect, the following example illustrates that this compound verb can be used as a non-locative alternation verb.

(40) John-ga himo-o sikkari-to musubi-age-ta.

John-NOM string-ACC firmly tie-up-PAST

'John tied up a string firmly.'

In (40), the verb *musubi-ageru* describes a resultant state of the theme *himo* 'string' brought about by the act of tying. As seen in (39), however, this compound verb cannot describe an event of moving the material to the box, nor can it describe the resultant state of the affected location *hako* 'box', which arises as a consequence of the act of tying a string. This suggests that the motion variant (39a) is excluded not because the complex verb *musubi-ageru* does not exist, but because the verb does not have a lexical specification for the change-of-state frame.

Incidentally, if *ageru* does not express a completive meaning, no restriction is imposed on the selection of the variants. This is illustrated by (41), where the second verb *ageru* indicates an upward motion.

(41) a. John-ga ude-ni hootai-o sita-kara maki-age-ta.

John-NOM arm-on bandage-ACC below-from roll-up-PAST

'John rolled up a bandage around his arm from below.' (Motion)

b. John-ga hootai-de ude-o sita-kara maki-age-ta.

John-NOM bandage-with arm-ACC below-from roll-up-PAST

'John rolled up the arm with a bandage from below.' (Change of State)

The data suggest that the verb *maki-ageru* 'wind up' can describe how the location (the arm) is affected by the act of rolling the bandage, as well as how the material (the bandage) is moved to the location. Since the oblique argument in (42b) cannot be deleted, we can confirm that *maku* 'roll' involves the perspective shift, which has the effect of adding the change-of-state frame to the verb *maku*, where it only has a lexical specification for the motion frame.

(42) a. John-ga (ude-ni) hootai-o mai-ta.

John-NOM arm-on bandage-ACC roll-PAST

'John rolled a bandage around his arm.' (Motion)

b. John-ga *(hootai-de) ude-o mai-ta

John-NOM bandage-with arm-ACC roll-PAST

'John rolled the arm with a bandage. (Change of State)

Thus, it should be clear that this type of compounding—i.e. compounding with the motional *ageru*—does not bring out the effect of excluding the syntactic frame provided

by the perspective shift.[15] In the light of this fact, we can see easily that for the purpose of checking whether a verb has a lexical specification for the change-of-state frame, it is necessary to look at the completive verb *ageru*, which expresses the meaning of 'completion' rather than 'motion'.

The data suggest that the presence or absence of the change-of-state variant involving the compound verbs with the completive *ageru* correlates with the question of whether or not the base verb obtains the change-of-state via the perspective shift.

(43)

	Oblique Argument Omission		V-V Compounding	
	Motion	Change-of-State	V-*tukeru*	V-*ageru*(completive)
musubu	√	*	√	*
sibaru/yuu	√	√	√	√

In the case of *musubu*, the change-of-state frame is derivationally created, which suggests that *musubu* has a lexical specification for the motion variant only. If the verb *ageru*, which has a completive meaning, picks out the change-of-state frame of the base verb, the unacceptability of (39b) follows, because *musubu* does not have a lexical specification for the change-of-state frame that the verb *ageru* should pick out.

Finally, recall that the verb *nuru* 'paint' allows the omission of the oblique argument in both motion and change-of-state variants, and that *nuru* can be compounded with either *tuku* 'attach' or *ageru* 'up'. With the compound verb *nuri-tukeru* 'paint-attach', only the motion variant is available, whereas only the change-of-state variant is possible with *nuri-ageru* 'paint-up', where the second verb *ageru* conveys a completive sense. The fact that the oblique argument can be dropped in the change-of-state variant of *nuri-ageru* suggests that this variant is not derived from the motion variant via the perspective shift.[16] The reverse also holds, since the oblique argument of *nuri-tukeru* in the motion variant can be dropped. The data show then that the verb *nuru* has a lexical specification for the two variants—i.e. both of the morphological forms are not derived via the perspective shift.

To summarize, in this section, by looking at compound verbs formed from *nuru* 'paint', *sibaru* 'bind' and *musubu* 'tie', we have argued that some locative alternation verbs should be ambiguous intrinsically, in that they have a lexical specification for both motion and change-of-state frames (i.e. the perspective shift does not provide any of these frames).

---

[15] The meaning of directionality is readily detected if *sita-kara* 'from below' is added.

[16] The verb *nuru* 'paint' can be combined with *ageru* 'up', which carries a completive meaning, but apparently, it does not take the material (i.e. the theme) as its accusative argument.

(i)   ?*John-ga   akai   penki-o   nuri-age-ta.
     John-NOM   red   paint-ACC   paint-up-PAST
     'John painted up the red paint.'

The compound verb implies that some complete change of state takes place on the object, but since the material *penki* 'paint' does not change its state merely by the act of painting, the sentence causes a conflict in meaning, hence, (i) should be excluded by a semantic deviance.

5. Conclusion

In this paper, mainly on the basis of compound verbs in Japanese, it has been argued that locative alternation verbs consist of the following two types: (1) one type which is lexically ambiguous in expressing both change-of-state and motion meanings (as its basic meanings), and (2) the other which has one of the two meanings as basic, while the other meaning is derived by the perspective shift (in the sense of Pinker (1989)). The Japanese data have shown that some locative verbs come to acquire a derivationally-created frame via the perspective shift. At the same time, the data suggest that other locative alternation verbs should be ambiguous intrinsically, in that they have a lexical specification for both motion and change-of-state frames (i.e. the perspective shift does not play a role in determining the possibility of the locative alternation).

These two types of verbs show distinct syntactic behavior. On the one hand, locative alternation verbs which are lexically ambiguous allow the omission of their arguments freely (insofar as their reference is recoverable from context). On the other hand, if one variant is derived via the perspective shift, it does not allow its oblique argument to be omitted. Accusative arguments are freely dropped in both basic and derived frames. The observed patterns of argument drop in Japanese show that for the perspective shift to be successful, the presence of a derivationally-created frame is syntactically identified by way of overtly expressing its oblique argument, while satisfying a certain semantic condition that allows us to recognize the meaning of an affected location.

References

Anderson, Stephen 1971, On the Role of Deep Structure in Semantic Interpretation. *Foundations of Language* 7, 387-396.

Dowty, David 1979, Dative "Movement" and Thomason's Extensions of Montague Grammar. In Steven Davis and Marianne Mithun (eds.) *Linguistics, Philosophy and Montague Grammar*, 153-222. Austin: University of Texas Press.

Fraser, Bruce 1971, A Note on the Spray Paint Cases. *Linguistic Inquiry* 2, 603-607.

Fillmore, Charles 1968, The Case for Case. In Emmon Bach and Robert Harms (eds.) *Universals in Linguistic Theory*, 1-88. New York: Holt, Reinhart and Winston.

Fukui, Naoki, Shigeru Miyagawa, and Carol Tenny 1985, Verb Classes in English and Japanese: A Case Study in the Interaction of Syntax, Morphology and Semantics. *Lexicon Project Working Papers* 3. Center for Cognitive Science, MIT.

Goldberg, Adele 1995, *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Huang, C.-T. James, and Y.-H. Audrey Li 1996, Recent Generative Studies in Chinese Syntax. In C.-T. James Huang and H.-Y. Audrey Li (ed.), *New Horizons in Chinese Linguistics*, 49-95. Dordrecht: Kluwer.

Iwata, Seizi 2008, *Locative Alternations*. Amsterdam: John Benjamins.

Jackendoff, Ray. 1990, *Semantic Structures*. Cambridge, Mass.: MIT Press.

Kageyama, Taro 1980, The Role of Thematic Relations in the Spray Paint Hypallage. *Papers in Japanese Linguistics* 7, 35-64.

Kishimoto, Hideki 2001, Locative Alternation in Japanese: A Case Study in the Interaction Between Syntax and Lexical Semantics. *Journal of Japanese Linguistics* 17, 59-81.

Levin, Beth 1986, The Formation of Adjectival Passive. *Linguistic Inquiry* 17, 623-661.

Levin, Beth 1993, *English Verb Classes and Alternations*. Chicago: University of Chicago Press.

Levin, Beth, and Tova Rapoport 1988, Lexical subordination. *CLS* 24, Part 1, 275-289.

Levin, Beth, and Malka Rappaport Hovav 2005, *Argument Realization*. Cambridge: Cambridge University Press.

Okutu, Keiichiro 1981, Idoo Henka Doosi Bun: Iwayuru Spray-Paint Hypallage-ni-tuite. *Kokugogaku* 127, 21-33.

Pesetsky, David 1995, *Zero Syntax: Experiencers and Cascades*. Cambridge, Mass.: MIT Press.

Pinker, Steven 1989, *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, Mass.: MIT Press.

Rappaport, Malka, and Beth Levin 1985, *A Case Study in Lexical Analysis: The Locative Alternation.* ms. Center for Cognitive Science, MIT.

Rappaport, Malka, and Beth Levin 1988, What to Do with θ-Roles. In Wendy Wilkins (ed.) *Syntax and Semantics 21: Thematic Relations*, 7-36. San Diego: Academic Press.

Tenny, Carol 1994, *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Theoretical and diachronic aspects of augmentation:
Evidence from Greek[1]

DIMITRA MELISSAROPOULOU

University of Patras

dmelissa@upatras.gr

IO MANOLESSOU

Academy of Athens

manolessou@academyofathens.gr

0. Abstract

This paper investigates the evolution of the morphological process of augmentative suffixation in the light of evidence provided by Greek. Two alternative pathways are examined: the evolution of augmentative suffixes from a) inflectional suffixes and b) derivational suffixes originally associated with different functions. On this basis, a novel theoretical proposal is elaborated, which assumes that augmentation is triggered by inflectional restructuring and gender changes for the differentiation of a) [+/- animate] and b) normal size vs. diminutive nouns. In this analysis, inflection seems to lie behind the formation of a new derivational process, that of augmentative suffixation, thus arguing against the discreteness of morphological processes and in favour of the existence of a morphological continuum, offering further support to the claim that there is considerable overlap between inflection and derivation.

1. Introduction

Evaluative morphology is a favourite topic in current linguistic theory; numerous papers have been published, focusing either on its position in grammar (among others Stump 1993, Bauer 1997, Scalise 1988, Melissaropoulou & Ralli 2008) or on the morphopragmatics and morphosemantics of the specific process (among others Jurafsky 1993, 1996, Dressler & Barbaresi 1994, Grandi 2002), with the onus of research lying on diminutives rather than augmentatives (however, cf. Dressler & Barbaresi 1994, Grandi 2002, Melissaropoulou forthcoming). This imbalance may perhaps be attributed to the fact that augmentatives display a rather limited productivity[2] compared to that of diminutives, something which might account for the fact that they have not equally drawn linguists' attention.

Augmentatives express the denotational meaning BIG, but may also realize a range of evaluative readings, such as exaggeration and intensification (Bakema & Geeraerts

---

[1] The authors would like to thank the audience of the 7th Mediterranean Morphology Meeting for useful comments and feedback. The first author also thanks the Greek State Scholarships Foundation for funding part of this work.

[2] Research on augmentatives is quite difficult to implement, because they belong to low registers and appear in texts very rarely (cf. Daltas 1985:63, Triantafyllides, 1991:125). This is a well known methodological problem with cross-linguistic validity.

2004:1045), and are generally viewed as the semantic counterpart of diminutives (cf. Schneider 1991 and Dressler & Barbaresi 1994).

The aim of this paper is to investigate the morphological process of augmentative suffixation in the light of evidence provided by the diachrony of Greek. More specifically, it hopes to shed light on the diachronic process of formation of augmentative suffixes. Two alternative pathways are examined: the evolution of augmentative suffixes from a) inflectional suffixes and b) derivational suffixes originally associated with different functions. Our data allow the elaboration of a new theoretical proposal, that the derivational process of augmentation can be triggered by inflectional restructuring and gender changes (cf. Corbett 1991, Grandi 2002).

Both functional and generative notions of the theory of morphological change (cf. Haspelmath's 1995 account of morphological reanalysis and Lass's 1997 notion of non-junk exaptation) will be evaluated against the data. The above discussion is intended as a contribution towards the central issue of the morphological status of augmentative formations as instances of inflectional or derivational processes, from the viewpoint of diachrony.

The paper is organized as follows: Section 2 provides the basic theoretical premises of the analysis. Section 3 sets out the data and describes the formal mechanism of affix creation. Section 4 develops the theoretical proposal. Finally, the conclusion summarizes the main points of the analysis.

2. Premises

Depending on the framework adopted, the mechanisms used to describe diachronic change can vary significantly and, in some cases, contradict each other.

One school of thought is the functionalist approach to *grammaticalization*, which is defined as a) a change whereby a unit acquires a grammatical or a more grammatical function it already has and/or b) a change strengthening the internal dependencies between the parts of a constructional schema (cf. Hopper & Traugott 1993, Bybee et al 1994, Heine et al 1991, Haspelmath 1998), a process which is standardly viewed as unidirectional[3]/[4]. The generative literature on the other hand tends to downplay the notion of grammaticalization (however, cf. Roberts 1993, Roberts & Roussou 2004), preferring instead to describe syntactic change on the basis of *reanalysis*[5].

Reanalysis usually refers to a change in the underlying structure that does not entail any change in the surface structure (see among others Langacker (1977:59)[6], Harris &

---

[3] This means that no development of less grammatical from more grammatical structures or elements is expected and that the attested counterexamples are relatively few or idiosyncratic, in some cases even disputable (cf. Traugott and Heine 1991, Haspelmath 1998, 2004, Kiparsky, 2005).

[4] For a different view cf. among others Ramat (1998), van de Auwera (2002).

[5] Different terms, except reanalysis, are used in the relevant literature to cover the same mechanism such as meta-analysis, back-formation, hyper-correction, folk-etymology etc.

[6] According to Langacker (1977:59) reanalysis is a *"change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation"*.

Campbell (1995:51)). Along with *extension* and *borrowing*, they are considered the only three mechanisms responsible for syntactic change (cf. Harris & Campbell 1995)[7].

A related term is that of *exaptation*, proposed by Lass (1997:316, 319), who defined it as "*a kind of conceptual renovation, as it were, of material that is already there, but either serving some other purpose, or serving no purpose. […] in exaptation, the 'model' itself is what's new*".

Crucially, considerable literature has been published discussing the interrelation of the above-mentioned mechanisms, ranging from the complete denial of the reanalysis mechanism (McDaniels 2003) or its dissociation from grammaticalization (cf. Haspelmath 1998[8]), to its inclusion or close relation to the latter (cf. Hopper & Traugott 1993[9], Heine et al (1991)[10] or even the reduction of grammaticalization to reanalysis (cf. Harris & Campbell 1995[11]).

Narrowing down these notions for the purposes of this paper, we should note that reanalysis cannot describe adequately the mechanisms involved in the evolution of Greek augmentative suffixation, since in the case of the -´_os and –a suffixes (see below for their description) it cannot predict the fact that their original function is maintained while a new one is simultaneously realized. Crucially, reanalysis involves change in the underlying structure, which in morphological terms corresponds to word or morpheme boundaries, something which does not occur with the morphemes in question. Unfortunately, the notion of reanalysis has been built and almost always exemplified on the basis of syntactic data, thus the implementation of this specific mechanism in morphology is not equally developed (but cf. Deutscher 2001) [12].

Additional terms necessary for the present analysis are the following adopted from Haspelmath's (1995) proposal concerning the growth of affixes in morphological reanalysis):

Affix-secretion (Haspelmath 1995:8), describing those cases where "*a non-affixal part of the root is reanalyzed as part of an affix*".

---

[7] For a different view arguing against the existence of reanalysis cf. McDaniels (2003:81-88). According to him a) reanalysis is not differentiated from exploratory expression and extension b) the role of pragmatics makes the use of the term reanalysis even more controversial and c) depending on the theoretical biases, (e.g. emergent grammar) reanalysis could be deprived of its reason of being.

[8] According to Haspelmath (1998: 315), "*The large majority of syntactic changes are instances of pure grammaticalization and should be explained within the framework of theory of grammaticalization without reference to reanalysis. […] Grammaticalization and reanalysis are disjoint classes of phenomena*".

[9] "*Unquestionably, reanalysis is the most important mechanism for grammaticalization*" Hopper & Traugott (1993:32).

[10] "*Typically reanalysis accompanies grammaticalization*" […] "*grammaticalization and reanalysis appear to be inseparable twins*" Heine et al (1991:27)

[11] As stated in Kiparsky (2005:4, 19, 31), Harris & Campbell (1995:90) reduced grammaticalization to reanalysis by distinguishing grammaticalization from analogy as "innovative reanalysis", in the sense that an existing category A is reparsed as a new category B.

[12] However, see Aronoff & Sridhar (1988:189) emphasizing on the role of reanalysis in morphological change, showing their antipathy for grammaticalization.

Affix-incretion (Haspelmath 1995:20), describing those cases where *"part of the affix is reanalyzed as part of the root"*.

Affix splitting (Haspelmath 1995:20), describing those cases where *"a single affix is split in two smaller ones that can be used separately"*.

Closing this section, what could be a preliminary observation is that contrary to Haspelmath's (1995) predictions, those patterns that are thought to be rare or counter-intuitive (i.e. *affix-incretion* and *affix splitting*) prove to play a crucial role for the evolution of a new derivational category, i.e. Greek augmentative suffixation.

3. Data: Description of the system

Let us start by stating the most obvious facts, which make Greek augmentative morphology conform to the patterns established on the basis of a considerable cross-linguistic sample (cf. Dressler & Barbaresi 1994, Grandi 2002, 2003[13]):

a) There are two types of suffixal augmentatives: Type I- those which augment the thing itself (*míti* 'nose' > *mitára* 'big nose') and Type II-those which indicate the person possessing the augmented thing (*míti* 'nose' > *mitarás* 'a person with a big nose')

b) Greek augmentatives are exclusively masculine and feminine, Greek diminutives belong to all three genders, but neuters are more frequent (Daltas 1985, Detsis, 1985, Melissaropoulou 2007, forthcoming).

c) Most complex augmentative suffixes developed from diminutive suffixes.

d) The implicational universal proposed by Schneider (1991) and Dressler & Barbaresi (1994), according to which the presence of augmentatives entails the presence of diminutives but not vice versa is valid for Greek as well. Greek developed suffixal augmentatives in the Medieval period (circa 12th c.) through the influence of diminutive morphology.

e) Greek seems to verify the well documented cross-linguistic trend towards iconic phonological realization of evaluative morphology,[14] in that augmentatives suffixes involve exclusively low and back vowels, while in diminutive suffixes the front vowel /i/ predominates.

We may group Modern Greek augmentative suffixes in 3 groups on the basis of their etymological provenance.

Group A: Bare inflectional suffixes: masculine: *-a, – ´_os*
(1) a. -a
maçér-a   <   maçéri
'big knife'   'knife'
varél-a   <   varéli

---

[13] Grandi (2003:140-141) attributes the evolution of the diminutive semantic value to the designation of a genealogical relation between the adult and the young, which he views as a 'strong typological matrix', acknowledging the lack of correspondence with the augmentatives. In his view, their evolution can be better accounted for in terms of areal constraints – tendencies.

[14] Cf. Wescott (1971)- despite the negative view expressed in Bauer (1996); the corpus he examines does show that the tendency exists at least in IE languages.

'big barrel'		'barel'
kolocíθa	<	kolocíθi
'big pumpkin'		'pumpkin'

   b. -_os

pórð-(os)	<	porðí
'big fart'		'fart'
tír-(os)	<	tirí
'big quantity/piece of cheese		'cheese'
mít-(os)	<	míti (Dod.)
'big nose'		'nose'

Group B: Complex derivational suffixes made up from (part of) a diminutive suffix plus a masculine or feminine inflectional marker: *-ara, -ar(os), -aka, -ak(os)*

(2) a. -ara

fon-ára	<	foní
'big voice'		'voice'
aft-ára	<	aftí
'big ear'		'ear'
poð-ára	<	póði
'big leg'		'leg'

   b. -ar(os)

péð-ar(os)	<	peðí
'big/handsome boy'		'child, boy'
koríts-ar(os)	<	korítsi
'big girl'		'girl'
scíl-ar(os)	<	scíl(os)
'big dog'		'dog'

   c. -aka (mainly in Mani)

velon-áka	<	velóni
'big needle'		'needle'
nis-áka	<	nisí
'big island'		'island'

   d. -ak(os) (Cyprus, Rhodes)

katsell-ák(os)	<	katsélla
'big chest'		'chest'
mastrap-ák(os)	<	mastrapá(s)
'big pot'		'pot'

Group C: Innovative augmentative suffixes: *-u/akla, -uklas, -u(m)ba*.

(3) a. -u/akla

psar-úkla	<	psári
'big fish'		'fish'
fon-ákla	<	foní

'big voice'		'voice'
mat-úkla	<	máti
'big eye'		'eye'

   b. -akla(s)

á(n)dr-akla(s)	<	á(n)dra(s)
'big/handsome man'		'man'
aráp-akla(s)	<	arápi(s)
'big negro'		'negro'

   c. -u(m)ba

kremið-ú(m)ba	<	kremíði
'big onion'		'onion'
cil-ú(m)ba	<	ciʎá
kafeð-ú(m)ba	<	kafes~kafeð
'big quantity of coffee'		'coffee'

Let us look at the historical process of evolution leading to the creation of the above groups of suffixes. Group A: the suffixes –a and –os were quasi- inflectional suffixes (cf. Haspelmath 1995: 3 for the term), marking specific inflectional classes (the "first" and the "second" declension, in traditional terms- cf. e.g. Woodard 2008: 24-26). In Ancient, Koine and Medieval Greek –a and –os were involved in inflectional and not in derivational processes.

These elements, while maintaining their original inflectional function, acquire an additional grammatical function. Within the framework of grammaticalization, this constitutes a special case of exaptation, namely non-junk exaptation (cf. Norde 2002: 55-56 who exemplifies via "a [Swedish nominal] suffix that was exapted for derivation, but retained its original inflectional function nevertheless"). Crucially, –a and –´_os still play an important role in Modern Greek nominal inflection (cf. Ralli 2000 for a detailed analysis).

According to traditional descriptions, the acquisition of augmentative meaning of these suffixes is mediated through the parallel existence of neuter diminutives, which had probably already lost their original diminutive meaning[15]. Schematically:

(4)

κεφαλ-ή > κεφάλ-ιον   > κεφάλ-ι(ν)   > κεφάλ-α
kefal-é: > cefál-ion   > cefál-i(n)   > cefál-a
'head'   'small head'   '(small) head'   'big head'

In the above example, no change in morpheme boundaries is observable. On the contrary, what has occurred is the reinterpretation of inflectional material as derivational one.

Passing now to group B suffixes, quite different mechanisms are in play. In both cases presented below, the point of origin seems to be derivational material, and more specifically diminutive suffixes. In the first case, part of the derivational suffix comes to

---

[15] cf. Jannaris (1897: 292) for relevant information.

be reanalysed as part of the root[16]. This, in terms of Haspelmath (1995) is *affix incretion*, a process which is rather counter-intuitive and thus expected to be rare. As a second step, the new reanalysed stem combines with the inflectional suffixes of group A, leading to a second reanalysis, *affix secretio*n this time: part of the root is reinterpreted as part of the affix, which creates a new complex suffix, able to act as an independent derivational element attaching to new roots. Schematically:

(5)

πούς > ποδ-άριον > ποδάρ-ιον   > ποδάρ-α > ποδ-άρα > φων-άρα   (< φωνή)
pu:s > poð-árion > poðár-ion   > poðár-a > poð-ára > fon-ára   (< foní)
'foot' 'small foot' '(small) foot'   'big foot' 'big foot'   'big voice' (< voice)

A more complex case is that of *–aka, -akos, -akas*. In this case, the diminutive suffix, which is itself the result of reanalysis (affix secretion again), is interpreted by the native speaker as being made up of two discrete constituents, because its final part, i.e. *-i*, is reanalyzed as inflectional marker. This is affix splitting in terms of Haspelmath (1995), again a counter-intuitive process. Schematically:

(6)

ρύακ-ς > ρυάκ-ιν   > ρυ-άκιν   > βελον-άκι   (< βελόνα) > βελον-ακ-ι
rýak-s > riák(c)-in   > ri-ácin   > velon-áci   (< velóna) > velon-ak(c)-i
'brook' 'small brook' '(small) 'brook 'small needle' (<needle)   > 'small needle''

               > βελον-άκ-ος
                 velon-ák-os
                 'big needle'

It would be difficult to argue that *-aki* became at any point the subject of affix incretion, namely that it was reanalyzed as part of the root. This is because it is by far the most productive diminutive suffix (cf. Melissaropoulou 2007 and references therein). Therefore, affix splitting is preferred over the analysis presented above for *–ari*, i.e. *affix incretion*, since the latter, from the early Middle Ages onward was no longer understood as a diminutive suffix.

Group C augmentative suffixes, i.e. *–u(m)ba, and -a/ukla, -a/uklas*, are of unclear etymological origin. *-a/ukla, -a/uklas*, according to Modern Greek dictionaries (Andriotis 2001, INS 1998, Babioniotis 1998) have an unclear connection with two homonymous latin suffixes *–c(u)lum* with a feminine ending *–a,* one bearing a diminutive meaning (e.g. *pisci-culus* 'small fish' < *piscis* 'fish'), and the other a deverbal one (e.g. *habitaculum* 'residence' < *habito* 'to inhabit'). The difficulty in attributing the Greek augmentative suffix to the Latin ones lies in the fact that the number of Latin loanwords in Greek with this ending is extremely limited (e.g. *panukla* < panucula (Lat.) 'plague', *kukla* < cuculla (Lat.) 'puppet') and their meaning has nothing to do with either

---

[16] The process is assisted by the borrowing of the latin suffix *–arium*, which has no diminutive meaning, and is extremely productive in creating neuter nouns (e.g. *armari* < *armarium* (Lat.)). Cf. Jannaris (1897: 293) and Minas (2003).

diminution or augmentation. As for *–u(m)ba*, it is not mentioned in Modern Greek dictionaries as it is a recent (21st century) formation, therefore no etymological information is available yet. The few older words with this ending are mainly of Latin or Turkish origin, bearing once again no evaluative meaning (e.g. *tulu(m)ba* 'a kind of sweet', *tu(m)ba* 'somersault'). Consequently, type C affixes will be excluded from the present discussion, since a consistent analysis is impossible at this point.

 Summing up, in this section, three basic mechanisms of augmentative suffix formation have been established. In the rest of the paper we will concentrate on the first two mechanisms, which are of considerable theoretical interest and represent two different paths of diachronic morphological change.

4. The proposed paths of evolution

In the cases described above, the parallelism with diminution was a prerequisite. This does not come as a surprise, since, as already mentioned, typological research has shown that augmentation entails the presence of diminution. However, diminutives constitute a necessary but not a sufficient condition for the evolution of their counterparts. Greek is a case in point: AG and Koine Greek possess a rich system of diminutive suffixation (-*ion, -idion, -arion* etc.) but only augmentatives of type II (i.e. those denoting a person bearing a property / characteristic in exaggeration). Augmentatives of type I appear only after the 12th c. AD.

Bearing this in mind, the triggering factor for the formation of augmentatives must be sought elsewhere, in some important morphological or semantic changes that took place around this period. One might postulate that the well established desemantization of several diminutive suffixes, such as *–ion* or *-arion*, might have been the initial causation for the acquisition of augmentative meaning. In the case of noun pairs, where one is the normal and its derivational pair is the diminutive, the loss of diminutive meaning in the second member and the consequent acquisition of 'normal' status would automatically push the meaning of the first member (the originally 'normal' one) by contrast to the meaning of an augmentative. This analysis, although plausible, runs up against the difficulty that desemantization is attested for diminutive suffixes from the 4th to the 10th century, whereas the new augmentatives begin to appear during the 12th century and are not well established before the 15th.

What we would like to suggest instead is that the evolution of augmentative suffixation is closely connected to the restructuring of nominal inflection, which does indeed take place during the period in question.

As a result of phonological changes leading to massive affix homophony, as well as extensive analogical levelling (cf. Schwyzer 1936, Seiler 1958, Ruge 1969, Holton & Manolessou 2010) the Ancient Greek nominal system, with its multiplicity of inflectional endings (cf. Sihler 1995, Woodard 2008) was considerably simplified, acquiring a closer connection with specific gender values. Whereas previously a noun's inflectional suffix was not predictable on the basis of its gender, from Medieval Greek onwards specific inflectional suffixes became extremely productive and as a result

characteristic of a particular gender and/or inflectional class (cf. Anastasiadi-Symeonidi & Cheila Markopoulou 2003, Christofidou 2003, in Anastasiadi-Symeonidi et al. 2003 for Modern Greek). This involves mainly 3 suffixes: *-a, -´_os and –i*.

For the feminine gender, the unmarked ending became –*a*, thanks to a) the massive transfer of old 3rd declension feminines to the 1st declension, *εἰκών > εἰκόνα, eikón > ikóna* 'picture' and b) the change of many feminines in *-i* to *-a, καλύβη > καλύβα kalyvi > kaliva* 'hovel'). A rough statistic of the first 5 volumes of Kriaras' lexicon of Medieval Greek (Kriaras 1967-) shows that about ¾ of feminine nouns end in –*a*. For the masculine gender, the unmarked inflectional ending became –*os* (however, cf. Christofidou 2003 for Modern Greek), thanks to the elimination of most non-masculine nouns ending in *-os*, either through transfer to the masculine gender (e.g. *ἡ ψῆφος > ὁ ψῆφος i psífos.MASC > o psífos.FEM* 'vote') or through replacement by their corresponding desemantized diminutive (e.g. *ἡ νῆσος > τὸ νησίν i ne:sos > to nisín* 'island')[17]. Lastly, for the neuter gender, –*i* became the unmarked inflectional ending, thus creating a new inflectional class, non existing in Ancient and Koine Greek, after having lost its diminutive meaning. It is true though that all neuter diminutives also end in –i (e.g. –*aki, -itsi, -uli*).

However, from the 12th c. onwards we observe the genesis and spread of non-neuter diminutive suffixes (e.g. –*pula, -itsa* etc.). This could serve as an indication that *-i* is no longer sufficient to denote diminution. Once *-i* acquires its new role of inflectional marker bearing the *par excellence* neuter gender value, it enters into an interchange relationship with the new unmarked inflectional endings for feminine and masculine, i.e. –*a* or –*os*. Change of gender in nouns, for whatever motivation, entails replacement of neuter –*i* by feminine –*a* and masculine –*os* endings. One could imagine two complementary contexts where this change would be required.

    a) In the case of animate neuter nouns, gender change would initially denote the corresponding feminine / or masculine of the species. According to Corbett (1991: 227-228) young animals which are not sex–differentiable are usually neuter, while the older and thus bigger of the species are masculine or feminine[18].

    b) In the case of neuter nouns in –*i* still retaining their original diminutive meaning, it would denote the normal sized de-diminutive noun.

These two contexts would create a multiplicity of noun pairs in the language whereby the second masculine or feminine is BIGGER THAN the first. These new pairs joined the already considerable number of preexisting noun pairs exhibiting this relationship, i.e. those pairs formed by nouns which originally bore the specific inflectional endings (-

---

[17] It is true of course that the masculine endings –*as* and –*is* were also available (the former especially since 3rd declension nouns acquire it, e.g. *χειμών > χειμώνας kheimo:n > çimónas* 'winter' and the latter forming productively deverbal agent nouns). However, roughly in the first five volumes of Kriaras lexicon, *-os* represents 65% of the masculine nouns, *-as / ás* 8%, while –*is / ís* 27%.

[18] "*Many indo-European languages assign sex- differentiable nouns to the masculine or feminine gender as appropriate, while the young of sex-differentiables – typically young animals which are treated as not yet sex-differentiable- are neuter*" Corbett (1991: 227-228)

os, e.g. *kláðos - klaðí* 'bigger branch'-'branch' *-a*, e.g. *ka(n)díla – ka(n)díli)* 'bigger cresset' – 'cresset' and which were not restricted to the categories a) and b) described above.

What is now in place is

(i) a mechanism which creates BIGGER THAN noun pairs through gender change in specific semantic contexts.

(ii) a large number of noun pairs which display this relationship without belonging to these semantic contexts.

Inevitably, analogy sets in, and extends the augmentative pattern to new neuter nouns. Once the gender interchange relationship extends to non animate and / or non diminutive neuter nouns in *–i*, this naturally leads to the reinterpretation of the second member as simply BIG.

Further support for this claim comes from the fact that neuter nouns of other inflectional classes cannot follow, at least not systematically, this derivational pattern. For example:

(7) címa   > *cimát-a.FEM
    'wave'   'big wave'
    vunó   > *vúna.FEM
    'mountain   'big mountain'

Once the replacement of *–i* by *–a* or *–os* is established as the semantic expression of the notion BIG, the new morphological category 'augmentative' is in place with two different suffixes: *-a* and *–os*[19]. However, this derivational pattern cannot apply to nouns that are not neuter, since they already bear the corresponding values, albeit with a different function, let alone the fact that the vast majority already bear the specific endings. The only way to augment these nouns is by creating new complex derivational suffixes, in which augmentative meaning is morphologically realized. The evolution described above can be schematized as follows:

Table 1:

1 >	2 >	3 >	4 >	5
Inflection: Variable	Inflection: -os, -a, -i	Inflection: -os, -a, -i	Inflection: -os, -a, -i	Inflection: -os, -a, -i
Augmentation: Ø	Augmentation: Ø	Gender change: a) animates b) normal-size Augmentation: Ø	Extension Augmentation: Type A for neuter nouns	Extension Augmentation: Type B and C for all nouns

---

[19] A factor not taken into account is the role of the accent. In fact, both originally inflectional suffixes *–a* and *–os* bear stress properties, in that they force the position of the accent of the noun to be in the penultimate and the antepenultimate syllable respectively –if possible-. These stress properties can be explained historically on the basis of analogy with preexisting *–i* vs. *–a* or *–os* pairs, in which the position of the accent was phonologically conditioned (e.g. *βουβάλι – βούβαλος vuváli - vúnalos* 'buffalo', *ελάφι – έλαφος eláfi - élafos* 'deer', *κοτσύφι – κόσσυφος kotsífi – kósifos* 'blackbird', *πλάτανι – πλάτανος platáni - plátanos* 'plane-tree'). On a synchronic level, these stress properties offer extra support to the derivational status of these elements and contrast with their homophonous inflectional counterparts.

Diminution: neuter	Diminution: -i	Diminution: variable	Diminution: variable	Diminution: variable

This evolutionary schema makes the following predictions:

a) The earliest innovative attestations of the –*os* and –*a* suffixes involve nouns which are not necessarily augmentative, but simply animate or normal-size.

(8) a. Animate nouns

ἀρκούδ-α *(Synax. Gad. 38)*	<	ἀρκούδ-ι(ν)
arkúð-a		arkúð-i(n)
'(female) bear'		'bear'
βουβάλ-α *(Ptochol. A. 119)*	<	βουβάλ-ι(ν)
vuvál-a		vuvál-i(n)
'(female) buffalo'		'buffalo'

    b. Normal size nouns

γούλ-α *(Ptoch. III 197-7 ms P app.)*	<	γουλ-ί(ν)
ɣúl-a		ɣúl-i(n)
'esophagus'		'(small) esophagus'
κουρούπ-α *(Landos, Geopon. 251)*	<	κουρούπ-ι(ν)
kurúp-a		kurúp-i(n)
'earthen pot'		'(small) earthen pot'

    c. Augmentative nouns

πιγούν-α (Spanos D 1658)	<	πιγούν-ι(ν)
piɣún-a		piɣún-i(n)
'big chin'		'chin'
παπούτσ-α (Limen., *Thanat.* 585)	<	παπούτσ-ι(ν)
papúts-a		papúts-i(n)
'big shoe'		'shoe'

b) The first attestations of complex augmentatives will appear attached not to neuter but only feminine and masculine bases.

(9)

μυτ-άρα	*(Synax. Gyn. 831)*	<	μύτη	+	-άρα
mit-ára			míti	+	-ára
'big nose'			'nose'.FEM		augment. suffix

φών-αρος *(Synax. Gad. 314)*		<	φωνή	+	-άρα
fón-aros			foní	+	-ára
'big voice'			'voice'.FEM		augment. suffix

c) Group A augmentatives should appear earlier in Medieval Greek than type B and C complex augmentatives. Although it cannot be verified on chronological terms (since they are all attested roughly in the same period), a strong indication for the chronological priority of type A suffixes is their greater productivity (more than 20 augmentatives of type A vs. only those presented above for types B and C).

5. Conclusions

In this paper we have discussed the evolution of a new morphological category, that of augmentation. We have elaborated a novel theoretical proposal, which assumes that augmentation is triggered by inflectional restructuring and gender changes for the differentiation of a) [+/- animates] and b) normal size vs. diminutive nouns.

Therefore, our facts and analysis do not verify Grandi's (2002, 2003) proposal for Greek, according to which possessive augmentatives, those of Type II, were the origin of pure augmentatives.

In our analysis, inflection seems to lie behind the formation of a new derivational process, that of augmentative suffixation, thus arguing against the discreteness of morphological processes and in favour of the existence of a morphological continuum in the spirit of Bybee (1985), offering further support to the claim that there is considerable overlap between inflection and derivation[20].

The specific morphological processes of formation of augmentative suffixes were described on the basis of different models, since different kinds of data can be handled more adequately by different theoretical models. More specifically, Haspelmath's (1995) reanalysis account provided the necessary terminology to describe the evolution of type B augmentatives, while Lass's (1997) notions of non-junk exaptation proved more appropriate for the description of type A augmentatives evolution.

The Greek data run counter to the predictions voiced in Haspelmath (1995), as affix-incretion and affix-splitting, two mechanisms considered as counter-intuitive and thus rare, were shown to have been the main factors behind the evolution of Greek augmentative suffixation. It should be noted that in Haspelmath's (1995) proposal, phonological reduction plays a key role in the interpretation of the morphological evolutionary pathways. However, this phenomenon is only marginally involved in the derivational process examined here. Consequently, one could suggest that these generalizations apply perhaps only to those languages where phonological reduction is the key factor.

Finally, the above proposed evolution for the formation of augmentatives in Greek, since it crucially involves discussion of changes in the inflectional system, indirectly sheds light on the native speaker's intuitions concerning inflectional properties and thus could serve as a basis for further research on nominal inflection.

References

Anastasiadi-Symeonidi, A. & D. Cheila Markopoulou 2003. Συγχρονικές και διαχρονικές τάσεις στο γένος της Ελληνικής (Μια θεωρητική πρόταση) [Synchronic and diachronic tendencies in Greek Gender (A theoretical account)]. In Anastasiadi-Symeonidi, A. Ralli, A. & D. Cheila-Markopoulou (eds.), *Γένος* [Gender]. Athens: Patakis, 13-56.

Anastasiadi-Symeonidi, A. Ralli, A. & D. Cheila-Markopoulou. 2003. *Γένος* [Gender]. Athens: Patakis.

---

[20] For discussion on the criteria distinguishing the two processes see among others Bybee (1985), Anderson (1992), Haspelmath (1996), Stump (1998)

Anderson, S. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Andriotis, N. 2001. *Ετυμολογικό Λεξικό της Κοινής Νέας Ελληνικής* [Etymological Dictionary of Standard Modern Greek], Thessaloniki: Institute for Modern Greek Studies. 3rd revised edition.

Aronoff, M. and S. N. Sridhar. 1987. Morphological Levels in English and Kannada. In Gussmann, E. (ed.), *Rules and the Lexicon*. Lublin: Catholic University, 9-22.

Babiniotis, G. 1998. *Λεξικό της Νέας Ελληνικής Γλώσσας* [Lexicon of Modern Greek]. Athens: Lexicolology Center.

Bakema, P. & D. Geerarts 2004. Diminution and Augmentation. In G. Booij, C. Lehmann, J. Mugdan & S. Skopeteas (eds.), *Morphology, An international Handbook on Inflection and Word-formation*. Berlin: Walter de Gruyter, 1045-1052.

Bauer, L. 1996. No phonetic iconicity in evaluative morphology. *Studia Linguistica* 50, 189-206.

Bauer, L. 1997. In search for Universals. *Studies in Language* 21, 3, 533-575.

Bybee, J. 1985. *Morphology. A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

Bybee, J. L., R. D. Perkins, and W. Pagliuca. 1994. *The evolution of grammar: Tense,aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Christofidou, A. 2003. Γένος και κλίση του ελληνικού ονόματος. Μια φυσική προσέγγιση [Gender and inflection of Greek nouns. A natural approach]. In Anastasiadi-Symeonidi, A. Ralli, A. & D. Cheila-Markopoulou. (eds). eds.), *Γένος* [Gender]. Athens: Patakis, 100-131.

Corbett, G. 1991. *Gender.* Cambridge: Cambridge University Press.

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.

Detsis, N. 1985. Ο υποκορισμός και η μεγέθυνση στη νεοελληνική [Diminution and augmentation in Modern Greek]. *Glossa 8*, 9-31.

Deutscher, G. 2001. On the mechanisms of morphological change. *Folia Linguistica Historica* 22:1-2, 41-48.

Dressler, W. & Barbaresi, M. 1994. *Morphopragmatics: Diminutives and Intensifiers in Italian, German and Other Languages* (Trends in Linguistics. Studies and monographs 76). Berlin/New York: de Gruyter.

Grandi, N. 2002. *Morfologie in contatto.* Milano: Francoangelli

Grandi, N. 2003. Matrici tipologiche vs. tendenze areali nel mutamento morfologico. La genesi della morfologia valutativa in prospettiva interlinguistica. *Lingue e Linguaggio* 3, 105-145.

Harris, A. & Campbell, L. 1995. *Historical Syntax in Cross-Linguistic Perspective.* Cambridge: CUP.

Haspelmath, M. 1995. The Growth of Affixes in Morphological Reanalysis. In G. Booij & J. Van Marle (eds.), *Yearbook of Morphology* 1994, 1-29. Dordrecht: Foris.

Haspelmath, M. 1998. Does grammaticalization need reanalysis?. *Studies in Language* 22:315-351.

Haspelmath, M. 2002. On directionality in language change with particular reference to grammaticalization. MS.

Haspelmath, M. 2004. On directionality in language change with particular reference to grammaticalization. In O. Fischer, M. Nurde & H. Norde (eds.), *The nature of grammaticalization.* Amsterdam: Benjamins, 17-44.

Heine, B. et al 1991. *Grammaticalization:a conceptual framework.* Chicago: University of Chicago Press.

Holton, D., & Manolessou, I. 2010. "Medieval Greek and Early Modern Greek". In E. Bakker (ed.), *A Companion to the Ancient Greek Language*, Oxford: Wiley/Blackwell, 539-563

Hopper, P. & E. Traugott 1993. *Grammaticalization.* Cambridge: CUP.

Institute for Modern Greek Studies [Manolis Triantafyllides Foundation]. 1998. *Λεξικό της Κοινής Νεοελληνικής* [Lexicon of Modern Greek Koine]. Aristotle University of Thessaloniki.

Jannaris, A. 1897. *An historical Greek grammar, chiefly of the Attic dialect.* London: MacMillan.

Joseph, B. 2001. Is there such a thing as grammaticalization?. *Language Sciences* 22, 163-186.

Jurafsky, D. 1993. Universals in the semantics of the diminutive. *Berkeley Linguistics Society* 19, 423 - 436.

Jurafsky, D. 1996. Universal Tendencies in the Semantics of the Diminutive. *Language* 72, 533-578.

Kiparsky, P. 1982. *Explanation in Phonology*. Dordrecht: Foris.

Kiparsky, P. 2005. Grammaticalization as optimization. Available on line at http://www.stanford.edu/~kiparsky/Papers/yalegrammaticalization.pdf. Accessed at 2010-04-13,

Kriaras, E. (ed.) 1967-. *Λεξικό της Μεσαιωνικής Ελληνικής Δημώδους Γραμματείας* [Lexicon of the Medieval Vernacular Literature]. Thessaloniki: Kentro Ellinikis Glossas, *1100-1669*.

Langacker, R. 1977. Syntactic reanalysis. In Li, Charles N. (ed.), *Mechanisms of syntactic change.* Austin: University of Texas Press, 59-139.

Lass, R. 1997. *Historical Linguistics and Language Change*. Cambridge: CUP.

Melissaropoulou, D. 2007. *Μορφολογική περιγραφή και ανάλυση του Μικρασιατικού ιδιώματος Κυδωνιών και Μοσχονησίων* [Word Formation in the Asia Minor Dialect of Kydonies and Moschonisia (Aivaliot)]. Unpublished PhD Dissertation. University of Patras.

Melissaropoulou, D. & A. Ralli. 2008. Headedness in diminutive formation: Evidence from Modern Greek and its dialectal variation. *Acta Linguistica Hungarica* 55: 183-204.

McDaniels, T. 2003. What's wrong with reanalysis. *Toronto Working Papers in Linguistics* 21:81-88.

Melissaropoulou, D. 2009. Augmentation vs. Diminution in Greek Dialectal Variation: an optimal system. In F. Montermini, G. Boyé, and J. Tseng (eds.), *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux.* Cascadilla Proceedings Project, Somerville, MA, USA, 125-137.

Minas, A. 2003. *Η μορφολογία της μεγέθυνσης στην ελληνική γλώσσα.* [The morphology of augmentation in the Greek language]. University of Ioannina, Research yearbook of The School of Humanities. Appendix No 8. Ioannina: Dodoni.

Norde, M. 2002. The final stages of grammaticalization. Affixhood and beyond. In I. Wischer & G. Diewald *New Reflections on Grammaticalization.* Amsterdam: Benjamins, 45-66.

Ralli, A. 2000. A Feature-based Analysis of Greek Nominal Inflection. *Glossolojia* 11-12: 201-228.

Ralli, A. Melissaropoulou, D. & A. Tsiamas 2004. Φαινόμενα αναδιάρθρωσης του ονοματικού κλιτικού παραδείγματος της διαλέκτου των Κυδωνιών και Μοσχονησίων [Restructuring Phenomena in the dialectal variety of Kydonies and Moschonisia]. *Studies in Greek Linguistics* 2003, 568-579.

Ramat, A. G. 1998. Testing the boundaries of grammaticalization. In A. G. Ramat & P. J. Hopper (eds.), *The limits of grammaticalization.* Amsterdam: Benjamins, 107-127.

Roberts, I. 1993. A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica* 13, 1-2: 219-258.

Roberts, I. & A. Roussou 2004. *Syntactic Change. A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge University Press.

Ruge, H. 1969: *Zur Entstehung der neugriechischen Substantiv-deklination*. Stockholm: Almquist & Wicksell.

Scalise, S. 1988. The Notion of Head in Morphology. In G. Booij & J. Van Marle (eds.), *Yearbook of Morphology* 1988. Dordrecht: Foris, 229-245.

Schneider, K. 1991. Affektive Lexik: Kognitive, semantische und morphologische Aspekte. In E. Klein, F. P. Duteil & K. H. Wagner (eds.), *Betriebslinguistik und Linguistikbetrieb: Akten des 24. Lingustischen Kolloquiums*, Universität Bremen 4.-6. September 1989. Tübingen: Niemeyer (Linguistische Arbeiten 260), 233-241.

Schwyzer, E. 1936. Zur Systematik der griechischen Nominaldeklination. *Glotta* 25 (1936) 205- 217.

Seiler, H. 1958. Zur Systematik und Entwicklungsgeschichte der griechischen Nominaldeklination. *Glotta* 37: 41-67.

Sihler, A. L. 1995. *New Comparative Grammar of Greek and Latin.* Oxford: Oxford University Press.

Stump, G. 1993. How peculiar is Evaluative Morphology? *Journal of Linguistics* 29, 1-36.

Stump, G. 1998. Comments on the paper by Inkelas and Orgun. In S. Lapointe, D. Brentari, and P. Farrell (eds.), *Morphology and its Relation to Phonology and Syntax.* Stanford: CSLI, 393–405.

Traugott, E. & B. Heine 1991. Introduction. In E. Traugott & B. Heine (eds.), *Approaches to Grammaticalization.* Amsterdam: Benjamins.

Triantafyllides, M. 1991. *Νεοελληνική Γραμματική* [Modern Greek Grammar]. 3rd edition with corrections. (1st edition 1941). Thessaloniki: Manolis Triantafyllides Foundation.

Van der Auwera, J. 2002. More thoughts on degrammaticalization. In I. Wischer & G. Diewald (eds.), *New reflections on grammaticalization,* 19-29. Amsterdam: Benjamins.

Woodard, R. (ed.) 2008. *The Ancient Languages of Europe*. Cambridge: CUP.

Wescott, R. W. 1971. Linguistic iconism. *Language* 47: 416-428.

Primary Sources

Synax. Gad. Συναξάριον τοῦ τιμημένου γαδάρου = Pochert, C. (1991): *Die Reimbildung in der spät- und postbyzantinischen Volksliteratur*. Köln: Romiosini.

Ptochol. A Πτωχολέων Α = Kechayoglou, G. (1978): *Κριτικὴ ἔκδοση τῆς Ἱστορίας Πτωχολέοντος. Θέματα ὑστεροβυζαντινῆς καὶ νεοελληνικῆς λογοτεχνίας*. Διδακτορικὴ διατριβὴ [*ΕΕΦΣΠΘ*, παράρτημα 22], Thessaloniki 1978.

Ptoch. III Στίχοι Θεοδώρου τοῦ Πτωχοπροδρόμου πρὸς τὸν βασιλέα κὺρ Μανουὴλ τὸν Κομνηνὸν = Eideneier, H. (1991). *Ptochoprodromos. Einführung, kritische Ausgabe, deutsche Übersetzung, Glossar*. Köln: Romiosini.

Landos, Geopon. Βιβλίον καλούμενον Γεωπονικόν ... = Kostoula, D. (1991) *Αγάπιος Λάνδος. Γεωπονικόν, Βενετία 1643*, Volos 1991.

Spanos D <Ἀκολουθία τοῦ Σπανοῦ> = Eideneier, H. (1977): Spanos. *Eine byzantinische Satire in der Form einer Parodie*. Berlin/NY: Walter de Gruyter & Co.

Synax. gyn. Συναξάριον τῶν εὐγενικῶν γυναικῶν = Krumbacher, K. (1905) "Ein vulgärgriechischer Weiberspiegel", *Sitzungsberichte der philosophisch-philologischen und der histor. Klasse der Königl. Bayer. Akademie der Wissenschaften*, Munich 1905, 335-433.

Limen., *Thanat. Τὸ θανατικὸν τῆς Ρόδου* = Legrand, E. (1880) "Ἐμμανουὴλ Γεωργιλλᾶ, *Τὸ θανατικὸν τῆς Ρόδου*", in *Bibliothèque grecque vulgaire*, vol. 1, Paris 1880, 203-225.

English [V-A]$_V$ forms and the interaction between morphology and syntax[*]

AKIKO NAGANO MASAHARU SHIMADA
University of Tsukuba *University of Tsukuba*
nagano.akiko.gt@u.tsukuba.ac.jp shimada.masaharu.fu@u.tsukuba.ac.jp

1. Introduction

In English resultative constructions, adjectival resultative predicates may occur not only in the post-object position, but sometimes also in the pre-object position. Compare the two sentences (a) and (b) paired in (1)-(4) below. The sentences in (a) are canonical resultative constructions, but in the corresponding sentences in (b), the adjective immediately follows the verb.

(1) a. Mother <u>bleached</u> the shirt <u>white</u>.

 b. Mother <u>bleached white</u> the shirt. (Taniwaki 2006:251)

(2) a. Mary <u>wiped</u> the floor <u>clean</u>.

 b. Mary <u>wiped clean</u> the floor. (Taniwaki 2006: 267)

(3) a. John <u>pushed</u> the door <u>open</u>.

 b. John <u>pushed open</u> the door. (Taniwaki 2006: 251)

(4) a. Cornelius <u>slammed</u> the boot <u>shut</u>.

 b. Cornelius <u>slammed shut</u> the boot. (Taniwaki 2006: 270)

As we will show immediately, previous studies have revealed that the V-A sequences in (1b)-(4b) behave as lexical units. However, it remains unclear what kinds of lexical units they are, and how they are related to the corresponding canonical resultative constructions in (1a)-(4a). The aim of this paper is to answer these questions and examine the implications they have for the interaction between morphology and syntax in English. Henceforth, we will call the V-A sequences "V-A forms."

The paper is organized as follows. In section 2, we will review previous studies on V-A forms and summarize the main properties of V-A forms. Their behavior as lexical units has led researchers to conclude that they are formed by morphological compounding, but as we will criticize in section 3, the compounding analysis has at least three serious problems. In section 4, we will offer an alternative analysis according to which V-A forms constitute a heterogeneous category. On the basis of new findings, we will argue that the V-A forms are not formed by compounding but arise from interactions between morphology and syntax. To be more specific, our claim is that the V-A forms in (1b) and (2b) are lexicalizations from the corresponding resultative constructions in (1a) and (2a), while the V-A forms in (3b) and (4b) and the corresponding resultative constructions in (3a) and (4a) represent particle verb constructions.

2. Previous Studies

As previous studies on V-A forms, we have Bolinger (1971), Kanemoto (2002), and Taniwaki (2006), among others. Although they differ in several respects, it seems safe to say that they agree that V-A forms have the following three basic properties:

(5)　　a. V-A forms are lexical units.

　　　　b. Semantically, V-A forms can be divided into the following two types:[1]

　　　 (i) The adjective is an overt realization of the resultant state that the verb inherently expresses.

　　　　　　e.g. (1b) *bleach white*, (2b) *wipe clean*

　　　　(ii) The verb expresses an activity of force exertion (e.g. *blow, fling, jerk, jiggle, pull, push, squeeze, throw, wave, wiggle*), and the adjective expresses the resultant state of disconnection (e.g. *clear, free, loose, open, shut*).

　　　　　　e.g. (3b) *push open*, (4b) *slam shut*

　　　　c. All the V-A forms have a corresponding resultative construction, but not vice versa.

To begin with the property stated in (5a), V-A forms are lexical units. Kanemoto (2002) and Taniwaki (2006) verify this by referring to standard criteria for the word-phrase distinction, including the Principle of Lexical Integrity (Bresnan and Mchombo 1995) and No Phrase Constraint (Botha 1984). Witness the data given in (6)-(8) below. The contrast in (6) shows that the gapping of the V-A form as a whole is possible, but deleting only the V is not allowed. According to the Principle of Lexical Integrity, this fact means that the V-A form as a whole constitutes a lexical unit. The lexical status of the V-A form is corroborated by the additional contrasts shown in (7) and (8); the adverbial modification of an adjective is possible in a resultative construction, but it is not in a V-A form. Also, it is possible to conjoin two V-A forms, but the conjunction of two adjectives inside a V-A form is not allowed. These observations are exactly what No Phrase Constraint predicts if the V-A form is a lexical unit.

(6)　　a. Eric <u>pushed open</u> the front door, and Tony [~~pushed open~~] the back door.

　　　　b. *Eric <u>pushed open</u> the front door, and Tony [~~pushed~~ shut] the back door.

　　　　　　　　　　　　　　　　　　　　　　　(Taniwaki 2006: 254)

(7)　　a. He <u>cut</u> the conference <u>very short</u>.

　　　　b. *He <u>cut very short</u> the conference. (cf. He <u>cut short</u> the conference.)

　　　　　　　　　　　　　　　　　　　　　　　(Kanemoto 2002: 83)

(8)　　a. John both <u>flung open</u> and <u>flung shut</u> the back door.

　　　　b. *John <u>flung</u> both <u>open and shut</u> the back door.　　　(Taniwaki 2006: 256)

　　In addition to the lexical status of V-A forms, Taniwaki (2006: 256-257) argues that the impossibility of stranding a PP complement of an AP, which is shown in (9b) below, means that V-A forms are lexical units formed in the lexicon rather than derived by incorporating the adjectival head of a resultative predicate into the main verb; that is, V-A forms are not instances of syntactic incorporation. Also, following Baker's (1985: 9) claim that "nominal gerunds" (*V-*

---

[1] In addition to these two types, Bolinger (1971: 72-73) also discusses the V-A form that consists of an "empty causative verb" and an adjective and expresses the existential sense "to present, reveal, bring on the scene." The examples are: *Have (make, hold, leave, keep) ready the answers.*/ *It renders necessary the measures.*/ *It makes plain the purpose.* We leave this type aside in this paper.

ing of NP) are formed in the lexicon, she says that the data given in (10) below provides positive evidence that V-A forms are formed in the lexicon.

(9)   a. He <u>wiped</u> the revolver <u>clean</u> of his fingerprints.

        b. *He <u>wiped clean</u>$_i$ the revolver t_i of his fingerprints.   cf. He <u>wiped clean</u> the revolver.

(10)   a. ??John's continuous <u>pushing</u> of the door <u>open</u> irritated his wife.

        b. John's continuous <u>pushing open</u> of the door irritated his wife.

Next, summarized in (5b) is the observation that V-A forms come in two semantic types. In the first type, the adjective in the V-A form overtly realizes the resultant state that the verb inherently expresses. This is the case in (1b) and (2b). For example, in (1b) *bleach white the shirt*, the verb *bleach* is an accomplishment verb that entails the resultant state "white," so the adjective *white* is a redundant expression of the verb's lexical meaning. In (2b) *wipe clean the table*, the verb is not an accomplishment verb, but an activity verb that denotes an activity with a specific purpose, which is in this case the purpose of making something clean. The adjective overtly realizes that presupposed purpose of the verb. In the words of Bolinger (1971: 74), the verb and adjective of this type of V-A form "represent some kind of cause-effect relationship in which the effect is more or less intrinsic to the cause."

The second semantic type of the V-A form is represented by (3b) and (4b). Basically, the verb is a "force exertion" verb like *blow, fling, jerk, jiggle, pull, push, squeeze, throw, wave*, and *wiggle*, and the adjective expresses disconnection or connection as in *clear, free, loose, open*, and *shut*. The example in (3a), *push open the door*, is a representative case of this type. The verb in (4a), *slam shut the boot*, differs from (3a) only in that it expresses force exertion accompanied by a sound emission event.

Finally, as stated in (5c), the V-A form always has a corresponding resultative construction, but not vice versa. This property of the V-A form is usually accounted for in terms of the semantic characterization discussed above. That is, it is argued that a resultative construction allows a V-A form only when either of the two semantic conditions in (5b) is met. Compare the data in (11) with the data in (12) below. In (11), we have both resultative constructions and V-A forms, but the resultatives in (12) do not have corresponding V-A forms.

(11)   a. John <u>drained dry</u> the glass.   cf. John <u>drained</u> the glass <u>dry</u>.

        b. They <u>cut short</u> the interview.   cf. They <u>cut</u> the interview <u>short</u>.

        c. He <u>whittled short</u> the stalk.   cf. He <u>whittled</u> the stalk <u>short</u>.

        d. He's {<u>planing/buffing/sanding</u>} <u>smooth</u> the boards.

             cf. He's {<u>planing/buffing/sanding</u>} the boards <u>smooth</u>.

        e. Mary <u>scrubbed clean</u> the floor.   cf. Mary <u>scrubbed</u> the floor <u>clean</u>.

        f. She <u>sucked dry</u> the orange.   cf. She <u>sucked</u> the orange <u>dry</u>.

        g. In the same instance he <u>flung open</u> the car door. cf. He <u>flung</u> the car door <u>open</u>.

        h. The old man <u>blew clear</u> a pipe.   cf. The old man <u>blew</u> a pipe <u>clear</u>.

        i. The prisoner <u>jerked free</u> his wrist.   cf. The prisoner <u>jerked</u> his wrist <u>free</u>.

        j. William <u>snapped shut</u> the lock.   cf. William <u>snapped</u> the lock <u>shut</u>.

        k. John <u>creaked open</u> the school gate. cf. John <u>creaked</u> the school gate <u>open</u>.

(12)   a. *Father <u>painted white</u> the fence.   cf. Father <u>painted</u> the fence <u>white</u>.

        b. *Mother <u>dyed black</u> her hair.   cf. Mother <u>dyed</u> her hair <u>black</u>.

        c. *She <u>shook awake</u> her husband.   cf. She <u>shook</u> her husband <u>awake</u>.

  d. *John <u>hammered flat</u> the metal.   cf. John <u>hammered</u> the metal <u>flat</u>.

  e. *The horses <u>dragged smooth</u> the logs.   cf. The horses <u>dragged</u> the logs <u>smooth</u>.

  f. *Stefan <u>ate clean</u> his plate.   cf. Stefan <u>ate</u> his plate <u>clean</u>.

  g. *He <u>danced sore</u> his feet.   cf. He <u>danced</u> his feet <u>sore</u>.

  h. *The chef <u>cooked black</u> the kitchen wall. cf. The chef <u>cooked</u> the kitchen walls <u>black</u>.

            ((11)-(12) from Taniwaki 2006: section 4.2)

The V-A forms in (11a-f) satisfy the semantic condition in (5bi), while the V-A forms in (11g-k) satisfy the semantic condition in (5bii). For instance, the accomplishment verb *drain* lexically entails the resultant state *dry*, so the V-A form *drain dry* is acceptable. And the V-A form *fling open* is allowed because it combines a force exertion verb and an adjective denoting disconnection. On the other hand, the V-A forms in (12) are unacceptable because they do not satisfy either of the semantic conditions in (5b). In particular, their verbs and adjectives are not in the relation of overt or redundant realization stated in (5bi). For instance, unlike the verb *bleach*, the accomplishment verb *paint* entails not a "white" state but a "colored" state, so unlike the form *bleach white*, the form *paint white* is unacceptable, as shown in (12a). The V-A forms *shake awake* and *hammer flat* in (12c, d) are unacceptable because their head verbs of an activity type have no intrinsically determined purpose to be realized by an adjective.

On the basis of these three observations in (5), Kanemoto (2002) and Taniwaki (2006) conclude that the V-A form is a compound verb; more strictly, Kanemoto calls it an "activity-result compound" and Taniwaki a "lexical compound." The compounding analysis will be critically examined and shown to be untenable in the next section. Bolinger (1971: chapter 6), on the other hand, proposes a view that that the V-A form is a particle verb, showing that the adjectives used in this form "are entitled to be classed with the particles that are used in phrasal verbs" (Bolinger ibid.: 71). In section 4, we will show that this view is correct for some of the V-A forms, but there exists other V-A forms that do not conform to the particle-verb analysis either.

3. [V-A]v forms as compound verbs

In this section, we will examine the implications of analyzing V-A forms as compound verbs and will demonstrate that such an analysis has at least three serious problems: the absence of verbal compounding in English, the systematic right-headedness of English endocentric compounding, and the non-homogeneous nature of V-A forms.

3.1. Verbal compounding in English

The phenomenon we are discussing is significant for the proper demarcation of English compounding. The property of V-A forms in (5a), that is, their status as "words," has led researchers to conclude that V-A forms are compound verbs. For example, Taniwaki (2006: 253) claims that V-A forms are "compound verbs formed in the lexicon," and the compounding is conditioned by the semantics. However, this analysis is in direct contradiction to the traditional view that verbal compounding is basically impossible in English. Witness the following quotations:

(13) a. In English, root compounds can be found consisting of combinations of the open
    categories N, V, and A;

    N N (*file cabinet*), N A (*sky blue*), A A (*icy cold*), A N (*hard hat*),

    <u>A V</u> (*dry farm*), <u>N V</u> (*handmake*), <u>V N</u> (*drawbridge*), <u>V V</u> (*stir-fry*)

Of these, only the first four types are productive, with N N compounds being by far the most productive compound type in English. Compounds containing V as one or both members are barely productive. (Lieber 2005: 378; underlines added)

b. [V]erbal compounds with nouns as non-heads are impossible in English, and […] verbs cannot incorporate adjectival/adverbial non-heads. For instance, neither *read a book*, *steal a car* nor *drive fast*, *move slowly* can be readily turned into compounds (**bookread*, **carsteal*, **fastdrive*, **slow(ly)-move*), whereas nominalized verbs and their arguments (as in *the reading of books*, *a driver of trains*) and deverbal adjectives and their adverbial/adjectival modifiers are happily condensed to compounds (*book-reading*, *train-driver*, *a fast-driving chauffeur*, *a slow-moving animal*). (Plag 2003: 154-155)

c. With the exception of verbs with preposed particles, verbal composition did not occur in Old English and does not seem to have existed in Germanic at all. […] Verbal composition does not exist in Present-day English either, though such verbs as *spotlight, blacklist, stagemanage* seem to contradict us. (Marchand 1969: 100-101)

According to Lieber (2005), of all the theoretically possible patterns of compounding shown in (13a), the underlined patterns that include V are unproductive. In (13b), Plag (2003) says that compound verbs with nominal, adjectival, or adverbial non-heads are impossible in English. In addition, as in (13c), Marchand (1969) argues that the absence of verbal compounding is a property of English as a Germanic language.

Notice that the verbs cited in (14) below might seem to be compound verbs as mentioned in (13c), but actually they are verbs derived by the three word-formation processes given in (15), that is, back-formation, conversion, and inversion. Therefore, the existence of verbs like those in (14) does not go against the descriptions we have seen in (13).

(14)　　a. N + V: to ghost-write, to head-hunt, to spoon-feed, to spotlight, to stage-manage

　　　　b. A + V: to blacklist, to cold-call, to dry-clean, to free-associate, to shortcut

　　　　c. V + V: to batter-fry, to stir-fry

　　　　d. Particle + V: to backfill, to download, to input, to upshift

(15)　　a. Back-formation　　stage-manager → to stage-manage, spoon-fed → to spoon-feed

　　　　b. Conversion　　spotlight$_N$ → to spotlight, blacklist$_N$ → to blacklist

　　　　c. Inversion　　load down → download$_{N/V}$, put in → input$_{N/V}$ (Berg 1998)

In sum, the first serious problem of the compounding analysis of the V-A form lies in the absence of verbal compounding from the system of compounding of English.

3.2. Right-hand Head Rule

The second problem of the compounding analysis is that English word-formation generally conforms to the Righthand Head Rule cited in (16) below. Except for a small number of category-changing prefixes like *de-* and *en-* (e.g. *debug, entomb*), English endocentric complex words are right-headed (cf. Booij 2005:78, Scalise 2008).

(16)　　Righthand Head Rule

　　　　In morphology, the head of a morphologically complex word is the righthand member of that word. (Williams 1981: 248)

Endocentric compounding in English also conforms to this rule; according to Lieber's (2009) comprehensive survey of the possible types of compounds in English, endocentric compounds

are always right-headed in this language. Then, if the compounding analysis were correct and the operating process were really compounding, the output compound should be not in a V-A form like *to bleach white*, but in an A-V form like **to white bleach*. However, as we can see in (17a) below, this right-headed form is unacceptable. Similarly, the examples in (17b) show that the right-headed versions of the V-A forms given in (2b)-(4b) are all unacceptable. In fact, A-V compounding seems to be more difficult than N-V compounding in English, for N-V compound verbs like *to truck-drive* are allowed if embedded in synthetic compounds, as shown in (18a) below, but A-V compound verbs are not allowed even under embedding, as shown in (18b).[2] That is, unlike N-V compound verbs, A-V compound verbs do not even have the "embedded productivity" (Booij 2009: 212-214).

(17)   a. *Mother <u>white-bleached</u> the shirt.   cf. (1b)

          b. *to clean-wipe, *to open-push, *to shut-slam   cf. (2-4b)

(18)   a. N+V: *to truck-drive, *to tax-pay   vs.   truck-driving, tax-payer

          b. A+V: *to flat-hammer   vs. *flat-hammered metal (Adam 2001: 94)

     As we will see in section 4.1, the V-A form is a left-headed lexical unit (e.g. *My mother {bleach<u>ed</u> white/*bleach whit<u>ed</u>} the shirt*). Proponents of the compounding analysis might claim that the V-A form is an exceptionally left-headed compound verb, but allowing for the left-headedness just for this type is obviously not a constructive analysis. It has no advantage except the viability of the compounding analysis, while it incurs the serious disadvantage of obliterating the systematic right-headedness of English endocentric compounding.

3.3. Classification

Thirdly, the compounding analysis treats V-A forms as a single unitary category, but the following observations suggest that such indiscriminative treatment is not appropriate:

(19)   a. The V-A form of the type in (5bi) — the *bleach white* type — is unproductive, and some native speakers do not accept its attested instances.

          b. The V-A form of the type in (5bii) — the *push open* type — is very productive and can be formed freely without referring to the corresponding resultative constructions. Its instances are consistently accepted by native speakers.

In section 2, we saw that V-A forms come in two semantic types; the V-A forms in (1b) and (2b) conform to the semantic condition in (5bi), while those in (3b) and (4b) observe the semantic condition in (5bii). These two semantic types differ also in productivity and acceptability. As stated in (19a), the V-A form of the semantic type in (5bi), which we will call "the *bleach white* type" for convenience, is unproductive, and native speakers' judgments on its instances are not consistent. That is, attested V-A forms of the (5bi) type can be rejected by native speakers. Hence, Bolinger (1971: 76-77) says that the *bleach white* type of V-A form is idiomatic and exhibits dialectal preferences and a sensitivity to the register. In contrast, as stated in (19b), the V-A form of the semantic type in (5bii), which will be called "the *push open* type" hereafter, is

---

[2] Notice that the following verbs have a meaning that refers to the corresponding compound noun (e.g. *to whitewash*: "to cover something with *whitewash*"), so they are not A-V compound verbs but derived verbs converted or back-formed from the compound noun:

(i)   to broadcast, to cold-call, to deep-freeze, to deep-fry, to double-check, to dry-clean, to dry-cure, to dry-fly, to dry-fry, to dry-shave, to dry-spin, to free-associate, to free-fall, to rough-cast, to rough-ride, to shortcut, to still-burn, to still-hunt, to whitewash (Nagano 2009)

very productive, and its examples are accepted consistently by native speakers. Taniwaki (2006: 268, 275) also notices the high productivity of the *push-open* type, mentioning that this type of V-A form can be "immediately formed and used in the appropriate context without referring to the corresponding resultative constructions." These differences in productivity and acceptability suggest that the *bleach-white* type and the *push-open* type should be treated as distinct linguistic forms.

The bifurcation of V-A forms into two different linguistic categories is also supported by the disjunctiveness of the semantic generalization of V-A forms given in (5b). For the *bleach white* type, the resultant state entailed by the verb is crucial, while in the *push open* type, the verb is an activity verb. Such a disjunctive generalization is very hard to reconcile with any unitary analysis of V-A forms, so it constitutes a serious problem not only for the compounding analysis but also for Bolinger's (1971: chapter 6) particle-verb analysis discussed at the end of section 2. In the next section, we will show that the *bleach white* type is a lexicalized verb, while the *push open* type is a particle verb.

4. [V-A]v forms as a heterogeneous category

In the previous section, we have revealed the following two things. First, V-A forms are lexical units, but they are not compounds. Second, V-A forms consist of (at least) two different types of lexical units. In this section, we will advance a new analysis that can account for not only these two findings but also other various properties of V-A forms. The basic insight we draw on here is that being a lexical unit does not necessarily mean being formed by a morphological process, and the lexicon can be expanded in non-morphological ways.

4.1. The *bleach white* type as a lexicalized verb

In section 3.3, we argued that the *bleach-white* type and *push-open* type of V-A forms need differentiated treatment. To begin with the first type, we would like to propose that V-A forms of the *bleach white* type arise from resultative constructions via the process of lexicalization. The term "lexicalization" has both synchronic and diachronic meanings. Synchronically, it refers to the listing of an item in the lexicon, while diachronically, it refers to phonological, semantic, or syntactic changes of an item (Hohenhaus 2005: section2; see also Brinton and Traugott 2005). Items to be lexicalized are most commonly complex words (e.g. Bauer 1983: chapter 3), but as the researchers cited below claim, syntactic phrases can also be lexicalized in both of the two senses.[3]

(20)   a. [P]hrases of fixed forms, after being generated above the X⁰-level in syntax, enter the lexicon and are listed as such in the lexicon, and they undergo reanalysis as lexical categories.   (Shimamura 2003: 643)

e.g. [P-the-N]ₚₚ, [P-a-N]ₚₚ: *over the fence gossip, in a row* nests

[V-the-N]ᵥₚ: *a connect the dots* puzzle

[N-P-a-N]ₙₚ: *a floor of a birdcage* taste   (ibid.: 632-633)

---

[3] The phrasal lexicalizations discussed in Sauer (2004: 1625-1628) and Fischer (2007: 80) (e.g. Old English *dōmes dæg* > Modern English *doomsday*, the impersonal clause *Me thinks that…* > the verbal form *methinks*) are of a diachronic nature, but in view of their idiosyncrasies, they must have been lexicalized also in the synchronic sense, i.e., they must have been listed in the lexicon.

> b. Words like *jack-in-the-box* are best regarded as lexicalized phrases, i.e. they are memorized holistically by the speakers. (Plag 2003: 136)
>
> c. There are complex items that function as words, yet whose internal structure is that of a clause or phrase rather than a compound.
>
>> e.g. NPs constituting phrasal words: *jack-in-the-box, stick-in-the-mud*
>>
>> an AP constituting a phrasal word: *dyed-in-the-wool*
>>
>> VPs constituting phrasal words: *couldn't-care-less, has-been, wannabe, forget-me-not* (Carstairs-McCarthy 2002: 67-69)

In particular, Shimamura (2003: 644), cited in (20a), argues that "the lexicalization of syntactic phrases of fixed forms can be accounted for by assuming that after such phrases are generated in syntax, they enter the lexicon, undergo the process of reanalysis, and are listed there as idioms which are X^0s." According to her, the underlined expressions in the compounds in (20a), such as *over the fence*, *connect the dots*, and *floor of a birdcage*, result from the reanalysis of a PP, VP, and NP as adjectives in the lexicon; more specifically, they are words created by applying reanalysis rules in the form of "Adj → [P–the–N]_{PP}" to syntactic phrases. The same author also shows in Shimamura (1986, 2000) that A-*to*-V compounds such as <u>*easy-to-understand instructions*</u>[4] and genitive compounds such as *woman's magazine* can each be analyzed as lexical units created by the lexicalization of syntactic phrases. Given the validity and prevalence of lexicalization as a way to produce a lexical unit from a syntactic unit, it would not be too far-fetched to hypothesize that a similar process is involved in our case as well. That is, we hypothesize that the V-A form of the *bleach-white* type arises when the resultative construction enters the lexicon and gets reanalyzed as a lexical category.

The lexicalization analysis can account for morphological and semantic properties of the *bleach-white* type as well as its low productivity. First, as the following examples show, an inflectional morpheme attaches not to the V-A form as a whole but to the V element. This fact means that V-A forms are perceived as left-headed lexical units. The left-headedness makes sense if the V-A forms are lexicalizations from syntactic phrases, which are left-headed in English.

(21) a. My mother {bleached white/*bleach whited} the shirt.

b. Mary {wiped clean/*wipe cleaned} the floor.

Carstairs-McCarthy (2002: 67) argues that the word status of *jack-in-the-box* (see (20b, c)) manifests itself in its plural form *jack-in-the-boxes*, where the plural morpheme attaches not to the head noun but to the whole expression. However, we have to notice that the internally inflected plural form *jacks-in-the-box* is also listed in dictionaries alongside the above form. Notice also that according to dictionaries, similar nominal expressions *jack-in-a-bottle* "long-tailed tit," *jack-in-a-box* "cuckoopint," *jack-in-office* "arrogant official," *jack-in-the-green* "participant in traditional May Day parades," and *jack-in-the-pulpit* "cuckoopint" all form their plural form by inflecting the head noun: *jacks-in-a-bottle*, *jacks-in-a-box*, *jacks-in-office*, *jacks-in-the-green*, and *jacks-in-the pulpit*. As long as these *jack*-expressions are best analyzed as lexicalized phrases, it must be the case that lexicalization does not (necessarily) nullify the internal structure of an input phrase.

---

[4] According to Shimamura (1986: 31-32), this type of lexicalized phrase exhibits certain variation in native speakers' acceptability judgments. Her informants all accepted *easy-to-understand instructions* and *a hard-to-master language*, but some of them were reluctant to accept *a comfortable-to-wear jacket*, *an interesting-to-read book*, and *a difficult-to-master language*. Notice that the *bleach-white* type of V-A form also exhibits variation in acceptability, as we discussed in (19a).

Second, the lexicalization analysis can account for the fact that adjectival passives based on resultative constructions are possible when the corresponding V-A forms are possible. According to Jackendoff (1990: 236), resultative phrases cannot in general be found in adjectival passives; he says that the adjectival passives given in (22a) below are either "at best marginal" or unacceptable. The instances in (22b) below are also rejected by Levin and Rappaport (1995: 43).

(22)   a. ??swept-clean room, ??squashed-flat grapes, *washed-clean clothes,

   *watered-flat tulips   (Jackendoff 1990: 236)

   b. *the run-thin pavement, *a ticked-awake baby, *a drunk-dry teapot

                                        (Levin and Rappaport 1995: 43)

However, there do exist some adjectival passives based on resultative constructions (Levin and Rappaport 1995: 43-44). Compare the examples given in (23a) below with those in (23b). Our informants rejected the adjectival passives in (23a) but accepted those in (23b). The latter two instances in (23b) and the example *the pounded-thin beef* are also accepted by Levin and Rappaport (ibid.).

(23)   a. *a painted-white wall, *a hammered-flat sword

   b. a bleached-white shirt, a wiped-clean table, pounded-flat metal

This acceptability difference neatly correlates with the existence of the V-A form; the adjectival passives in (23b) have corresponding V-A forms (e.g. *to bleach white, to wipe clean*), while those in (22) and (23a) do not (*to paint white*, *to run thin*). That is, this type of adjectival passive cannot be formed without a V-A form listed in the lexicon. For instance, *a bleached-white shirt* is possible because the verb *to bleach-white* is listed in the lexicon, while *a run-thin pavement* is impossible because we do not have the verb *to run-thin* in the lexicon.

The strongest support for the lexicalization analysis comes from the semantic properties of the *bleach-white* type. Unlike the *push-open* type, this type of V-A form is not synonymous with the corresponding resultative constructions. Specifically, the *bleach-white* type can express an idiomatic meaning absent from the resultative counterpart and exhibit a semantic change that can be called "anti-resultativization." To begin with the idiomaticity, the semantic contrast shown in (24) below provides a simple but clear demonstration. While the resultative construction *to cut something short* expresses the compositional meaning "to make something short by cutting," the lexicalized form *to cut short something* has an idiomatic meaning "to bring something to a sudden end." The latter meaning cannot be expressed by the resultative.

(24)   a. to cut …. short "to make short by cutting"

      e.g. I can't decide whether or not to cut my hair short.

   b. to cut short "to bring to a sudden end" (*OED*)

      e.g. Death suddenly cut short his promising life.

In addition to the idiomaticity, the *bleach-white* type of V-A form differs from the corresponding resultative construction in that it does not entail a resultant state, expressing only an activity with a specific purpose. Witness the contrast shown in (25a, b) below. A resultative construction cannot be canceled, but the cancellation is possible in the *bleach-white* type of V-A form. Importantly, the *push-open* type of V-A form is similar to the corresponding resultative construction in that it disallows cancellation, as shown in (26).

(25)   a. *John bleached the shirt white, but the stain remained.

   b. John bleached white the shirt, but the stain remained.

      cf. *John hammered the metal flat, but it didn't become flat.

(26)   a. *John <u>pushed</u> the heavy door <u>open</u>, but it didn't move.

        b. *John <u>pushed open</u> the heavy door open, but it didn't move.

This observation suggests that while the *push-open* type is aspectually similar to the corresponding resultative construction, the *bleach-white* type is not; it is not an accomplishment verb but an activity verb.

Additional data based on Dowty's (1979) criteria for the aspectual classification confirm this conjecture. According to Dowty (1979), activity verbs occur with a time adverbial phrase headed by *for*, while accomplishment and achievement verbs take a time adverbial phrase headed by *in*, as shown in (27) below. Similarly, as show in (28) below, activity verbs differ from accomplishment verbs in that they cannot occur with the verb *finish*.

(27)   a. John walked for an hour./(*) John walked in an hour.

        b. ?John painted a picture for an hour./John painted a picture in an hour.

        c. ??John noticed the painting for a few minutes. /John noticed the painting in a few

            minutes.   (Dowty 1979: 56-58)

(28)   a. John finished painting a picture.

        b. *John finished walking.   (Dowty 1979: 57-59)

Witness our informants' judgments shown in (29) and (30) below. As shown in (29) and (30a, b), the *bleach-white* type of V-A form takes a *for*-phrase rather than an *in*-phrase and cannot occur with *finish*, so it must be an activity verb. This means that unlike the resultative phrase *bleach the shirt white*, the lexicalized verb *to bleach white the shirt* means to "do a certain type of washing activity to the shirt in order to make it white," without implying the accomplishment of the purpose. On the other hand, the *push-open* type of V-A form is an accomplishment verb and entails a resultant state, as shown in (30c, d).

(29)   a. John <u>bleached white</u> the shirt for an hour.

        b. ??John <u>bleached white</u> the shirt in an hour.

(30)   a. John finished <u>bleaching</u> the shirt <u>white</u>.

        b. *John finished <u>bleaching white</u> the shirt.

        c. John finished <u>pushing</u> the heavy door <u>open</u>.

        d. John finished <u>pushing open</u> the heavy door.

    To summarize the discussion so far, the *push-open* type of V–A form is basically synonymous with its separate form, but the *bleach-white* type of V–A form is peculiar in its idiomatic meaning and "anti-resultative" semantics. These peculiarities can be attributed to the lexicalized status of the *bleach-white* type. It is generally observed that "words, when embedded in complex words, lose their referential potential (in fact, it is not words but phrases that refer to something)" (Booij 2005: 188). Then, it must be the case that when the V-A sequence as a whole is reanalyzed as a word, the adjective within loses its referential potential and cannot express an independent stative event.[5] If so, the semantic difference between the V-A form *bleach white the shirt* and the resultative *bleach the shirt white* is similar to the semantic difference between the

---

[5] Closely related to the loss of referential potential is the observation that nouns inside words are interpreted in a generic sense and do not refer specifically to particular entities. For example, *lion* in the compound noun *lion-hunter* refers to the class "lion" generally. Shimamura (2003: 635) shows that nouns in lexicalized phrases also have generic interpretations, hence the following contrast: *an after-the-party mess* vs. *an after-the-party-given-by-Bill mess*. Our assumption here is that a similar change of interpretation should occur also in adjectives when they are put inside words.

derived verb *shorten* and the verb phrase *make (something) short*. For instance, *to shorten the skirt* means to "make the skirt shorter" rather than to "make the skirt short"; that is, the adjective incorporated into the word is no longer referential and so cannot denote a specific type of shortness.

Finally, the low productivity of the *bleach-white* type and its variability of native speakers' judgments, which we saw in (19a), can also be accounted for by the lexicalization analysis. First of all, the lexicalization process, whether it is synchronic or diachronic, is not an obligatory process that applies to all the relevant forms. Shimamura's (2003) reanalysis rules (e.g. Adj → [P-the-N]$_{PP}$), for example, do not apply to all the syntactic phrases with the specified structure. In addition, being a word, a lexicalized phrase should observe the nameability requirement, a pragmatic requirement imposed on words (Bauer 1983: 86-87). Nameability of a concept crucially depends on how well the concept is established in the language society. For instance, Kato and Kageyama (1998: 314) point out that in order for a [XP-N] form to achieve the lexical status as a phrasal compound, the reanalysis of the modifier XP along the lines of Shimamura is not sufficient; they argue that it is necessary that "a whole modifier-head unit (e.g. *inside-the-park homerun*) should be established as a conventional concept in the language society."

Similarly, in our case, a resultative construction can be lexicalized as a V-A form only when the activity associated with it is established as a conventional activity. If this view is on the right track, the lexicalization analysis can explain that the following V-A forms, mostly repeated from (12), are unacceptable because they do not fulfill the nameability requirement:

(31)   a. Weak resultatives: *to <u>paint white</u> the fence, *to <u>dye black</u> the hair,

        *to <u>shake awake</u> husband, *to <u>hammer flat</u> the metal

      b. Strong resultatives: *to <u>dance sore</u> one's feet, *to <u>cook black</u> the kitchen wall,

        *to <u>drink dry</u> the teapot (vs. to drink the teapot dry),

        *to <u>run thin</u> the pavement (vs. to run the pavement thin)

A lexicalized V-A form expresses to "do a certain activity in order to bring about a certain state," but such activities associated with these examples are not perceived as conventional enough to deserve naming. Take the V-A form **to paint white* given in (31a), for instance. Kanemoto (2002: 93) and Taniwaki (2006: 266-267) attribute the unacceptability of this form to the lack of the relation of overt semantic realization between V and A (see (5bi)), but in our view, this V-A form could become acceptable if the activity it denotes is established as a convention for one reason or another. For example, it could be accepted by members of a theatrical circle where doing a certain type of painting to an actor's face to make it white is established as a conventional procedure. Notice that not only the so-called strong resultatives but also weak resultatives (Washio 1997) cannot be lexicalized unless they attain the nameability. The factor that distinguishes between acceptable and unacceptable V-A forms is not the weak vs. strong distinction of the corresponding resultative construction, as implied in Taniwaki (2006: section 4), but the extralinguistic appropriateness of the naming.[6]

A similar claim is made by Bolinger (1971: 76), who argues, as we mentioned in section 3.3, that "idiomaticity" is a factor that makes many theoretically possible V-A forms of the *bleach-white* type unacceptable. According to him, the following contrast is due to the extralinguistic fact that extension is normal for width but not for length:

---

[6] Hence, acceptable V-A forms of the *bleach-white* type come from both weak and strong resultatives: e.g. Weak: *to bleach white (the shirt)*, Strong: *to pound soft (the clay)* (Bolinger 1971: 75).

(32)　　a.　They stretched (pulled, spread, drew) wide the fabric.

　　　　b. *They stretched (pulled, drew) long the rope.　　　(Bolinger 1971: 76)

To put this claim differently, the V-A forms in (32a) is acceptable because it denotes a conventionalized activity, but the activity denoted by the V-A forms in (32b) is not conventional and so is not nameable.

Given this extralinguistic property of lexicalization (see also Lipka 1992),[7] it is only natural that V-A forms of the *bleach-white* type cannot be formed freely and can be judged differently among native speakers.[8]

4.2. The *push open* type as a particle verb

As we mentioned in section 2, Bolinger (1971: chapter 6) claims that particles used in the so-called phrasal verbs are not restricted to the class "Adprep" (adverbs that function also as prepositions; e.g. *away, in, out, over*) and proposes a hypothesis that the V-A form belongs to the phrasal verb. In this section, we will show that this hypothesis is correct as long as the *push-open* type of V-A form is concerned.

In the preceding section, we have seen that unlike the *bleach-white* type, the *push-open* type of V-A form is basically synonymous with the corresponding separate form; neither of the forms allows cancellation, as in (26), and both of them can occur with the verb *finish*, as in (30c, d). This semantic synonymy and the high productivity we saw in (19b) can be accounted for if the *push-open* type of V-A form and its separate form are realizations of the same particle verb construction. That is, we propose that the V-A form *to push open the door* and its separate form *to push the door open* correspond respectively to the adjacent and separate forms of a canonical particle verb combination exemplified below.

(33)　　a. John <u>pulled off</u> the leeches.

　　　　b. John <u>pulled</u> the leeches <u>off</u>.

Particle verb constructions such as the one in (33) are generally considered to have a hybrid character, the adjacent form in (33a) behaving as a lexical unit and the separate form in (33b) behaving as a syntactic combination. There exists a considerable amount of literature on this hybrid character of the particle verb construction (see, for instance, Dehé et al. 2002, Spencer 2005: 79-81, among others), and we cannot go into a detailed examination of the relevant

---

[7] Lipka (1992: 7-8) cites the high frequency of use as a necessary condition for lexicalization. Also he says that lexicalization depends on "different regional, social, stylistic and other varieties of a language."

[8] Although we cannot go into details in this paper, we should clarify conditions for lexicalization not only from semantic and pragmatic points of view, but also from a structural viewpoint. The existence of structural constraints on the lexicalization of a V-A form is indicated by the fact that a *depictive* type of V-A form was totally unacceptable for any of our informants:

(i) *John <u>ate raw</u> the fish. (vs. John <u>ate</u> the fish <u>raw</u>.)

The following data indicate that a resultative predicate is structurally closer to the main verb than a depictive predicate:

(ii) John hammered the metal <u>flat</u> <u>hot</u>.　　　　　　(cf. McNulty 1988: 38)

The depictive predicate *hot* cannot precede the resultative predicate *flat* in this sentence. Also, the *do-so* test demonstrates that a resultative predicate forms the smallest constituent with the verb, but a depictive predicate does not:

(iii)　　a. John ate meat raw, and Tom did so rare.

　　　　b. *John painted a house red, and then Tom did so blue.　　　(Hoshi 1992: 9)

previous studies. One thing that we are convinced of, however, is that the adjacent Verb-Particle form such as *to pull off (the leeches)* cannot be formed morphologically in English; they cannot be analyzed as compound verbs for the same reasons that we offered in section 2 as objections against the compounding analysis of the V-A form (see also Zeller 2002: 255-256). English morphology does not allow compound verb formation, and English endocentric compounds are right-headed. If the compounding analysis for the particle verb were correct, the output should be *to off-pull* rather than *to pull off*, but the Particle-Verb sequence is generally not allowed in English (see Berg 1998, Ackema and Neeleman 2004: 159-160).[9] Then, we need to treat the Verb-Particle form as a lexical unit formed non-morphologically. In fact, not a few researchers propose theories of the particle-verb construction that are consistent with this status of the Verb-Particle form. For instance, Booij (2002) proposes that particle verbs are "constructional idioms" in the form "[X [ ]ᵥ]ᵥ' where X = P, Adv, A or N" that are created in the lexicon. He says that this is "the formation in the lexicon of units that are functionally identical to complex words, but do not form one grammatical word, but two" (Booij 2002: 40). The status of the Verb-Particle form as a non-morphological lexical unit can also be accounted for by Zeller's (2002) syntactic approach, which claims that particle verbs basically have a VP structure, but they can also have a V^0 structure when the phrasal structure is reanalyzed as a complex head.

Let us proceed to the particle-verb analysis of the *push-open* type of V-A form. Our claim is that the *push-open* type acquires its lexical status in the same way as the adjacent form of a particle verb construction does so (see above). To begin with, the right-headed form of this type (e.g. *to open push*; see (17b)) is unacceptable because the Particle-Verb form is unacceptable in English, as we have discussed just above. Secondly, as we saw in (5bii), the *push-open* type consists of a force exertion verb such as *jerk*, *pull*, *push*, and *throw*, and one of the disconnection adjectives in the set {*clear, free, loose, open, shut*}. These adjectives qualify as particles in that they form a closed class and express a change of location; Bolinger (1971: 85) defines the particle semantically as follows: "the particle must contain two [semantic] features, one of motion-through-location, the other of terminus or result." The disconnection adjectives can be conjoined with adverbial particles, as in *With a bound he was away and free* (Bolinger 1971: 68). Next, consider the following data concerning the morphosyntactic properties of the *push-open* type and particle verbs:

(34)   a. John <u>pulled</u> {<u>loose</u> the leeches / the leeches <u>loose</u>} and Tom the seaweed.

b. John <u>pulled</u> {<u>off</u> the leeches / the leeches <u>off</u>} and Tom the seaweed.

(35)   a. John <u>pushed</u> {*wide <u>open</u> the door / the door wide <u>open</u>}. (Taniwaki 2006: 255)

b. Fran <u>put</u> {*right <u>together</u> the model airplane / put the model airplane right <u>together</u>}.

(Jackendoff 2002: 71)

(36)   a. ??John's continuous <u>pushing</u> of the door <u>open</u> irritated his wife. [= (10a)]

b. John's continuous <u>pushing open</u> of the door irritated his wife. [= (10b)]

c. *The rapid <u>looking</u> of the information <u>up</u> is important.

d. The rapid <u>looking up</u> of the information is important.   (Jackendoff 2002: 72)

---

[9] According to Berg (1998: section 4), the difficulty of the Particle-Verb sequence in English arises from its word order pattern of SVO. As cautioned in Ackema and Neeleman (2004: 160), verbs such as *outperform*, *overact*, and *underfeed* are not related to the particle verb construction (e.g. *to perform out*, *to act over*, *to feed under*). They are verbs derived by prefixation.

These data show that the *push-open* type exhibits the same morphosyntactic properties as particle verbs. The sentences in (34) attest to the equal possibility of gapping, and those in (35) and (36) speak for the lexical status of their adjacent forms; the adjacent forms of the *push-open* type and particle verb both observe the Principle of Lexical Integrity, and they permit the nominalization by the derivational suffix *-ing*, or the formation of the "nominal gerund," in contrast to the separate forms. In addition to this parallelism, our claim that one should distinguish the *push-open* type of V-A form and particle verbs from the *bleach-white* type of V-A form is confirmed by the contrast between the derivatives in (37a, b) and those in (37c) given below.

(37) a. passer by, come outer, cleaner upper   (Ackema and Neeleman 2004: 160-161)

washer-up, let-downer, washer-upper, clean-uppable   (*WebCorp*)

b. pusher-open, push-opener, a real slam-shutter, pull-openable

c. ??bleach-whitable, *bleachable-white, ??a cut-shortable tutorial

It is well known that the position of a derivational suffix attached to a particle verb is variable, which is shown in (37a). Our informants' data given in (37b) show that the *push-open* type exhibits the same positional variability of a derivational suffix. Similar derivatives can be found also in the *WebCorp*, a corpus on the Internet (see [http://www.webcorp.org.uk/]). The *bleach-white* type of V-A form, on the other hand, strongly resists any further derivation. Derivatives such as those given in (37c) were never accepted by our informants nor were rarely found in the *WebCorp*. Notice that this is another similarity between the *bleach-white* type of V-A form and lexicalized phrases in general; the lexicalized phrases cited in (20) also resist undergoing derivation (e.g. *over-the-fenceness*, *jack-in-the-boxish*, *dyded-in-the-woolness*, ?*has-beenish*).

Fourth, the formal separability of particle verbs as well as the *push-open* type does not affect their aspectual property. Consider the following sentences:

(38) a. *John <u>pulled</u> the leeches <u>off</u>, but they still stuck to him.

b.*John <u>pulled off</u> the leeches, but they still stuck to him.

c. *John <u>pushed</u> the heavy door <u>open</u>, but it didn't move.   [= (26a)]

d. *John <u>pushed open</u> the heavy door open, but it didn't move.   [= (26b)]

e. *John <u>bleached</u> the shirt <u>white</u>, but the stain remained.   [= (25a)]

f. John <u>bleached white</u> the shirt, but the stain remained.   [= (25b)]

(39) a. John finished <u>pulling</u> the leech <u>off</u>.

b. John finished <u>pulling off</u> the leech.

c. John finished <u>pushing</u> the heavy door <u>open</u>.   [= (30c)]

d. John finished <u>pushing open</u> the heavy door.   [= (30d)]

e. John finished <u>bleaching</u> the shirt <u>white</u>.   [= (30a)]

f. *John finished <u>bleaching white</u> the shirt.   [= (30b)]

As in (38a, b), the adjacent and separate forms of a particle verb both refuse cancellation, and as in (39a, b), they both allow the co-occurrence with the verb *finish*. This fact means that the adjacent and separate forms of a particle have the same aspectual property. As the additional sentences given in (38c-f) and (39c-f), repeated from section 4.1, show, this aspectual stability is shared by the *push-open* type of V-A form, while the *bleach-white* type of V-A form undergoes the anti-resultativization semantically. Bolinger (1971: 82) claims that the adjacent and separate forms of a particle verb are semantically different in that "though the phrasal verb embodies both the action and the result, the position of the particle tends to make one or the other

paramount." According to him, the preposed particle makes the action paramount, whereas the postposed particle makes the result paramount. However, the on-going discussion shows that this difference is no more than a difference in focus and does not affect the aspectual property of a particle verb.

Lastly, the *push-open* type of V-A form and its separate form are similar to the two forms of a particle verb construction in that both forms are accepted and used equally freely. Witness the following data, where we compare three native speakers' judgments on the *bleach-white* type of V-A form, the *push-open* type of V-A form, and the adjacent form of a particle verb:

(40)

	Bleach white type	*Push open* type	*Pull off* type
a Canadian speaker	OK in some cases	OK	OK
an American speaker	OK in passive	OK	OK
a British speaker	unacceptable	OK	OK

cf. The separate forms (V…A/P) were accepted in all of the three types.

The three informants are all university-level English teachers and come from Canada, America, and Britain respectively. After we confirmed that they accept the separate V…A/P forms (e.g. *Mary bleached the shirt white/ Bill pushed the door open/ Tom pulled the leech off*), we asked them to judge the acceptability of the adjacent counterparts. As we can see, their judgments on the *bleach-white* type of V-A form were inconsistent, but they accepted the other two types consistently and without any hesitation. To be more specific with the results of the *bleach-white* type, the Canadian speaker accepted some of its instances (e.g. *Mary bleached white the shirt.*) but rejected others (e.g. *Kill dead the cockroach!*). The American speaker mentioned that the instances are permissible if used in passive, while the British speaker rejected all the instances of this type of V-A form. Such inconsistencies among native speakers were not observed in the other two types of adjacent forms, which were always accepted by all of the three informants. It is also significant to note that examples of the V-A form based on a depictive construction were totally and consistently rejected by all of the three informants (e.g. **John ate raw the fish* vs. *John ate the fish raw*). Compared with the crystal-clear rejection observed in this case, the inconsistent acceptability of the *bleach-white* type should be attributed not to some structural factors but to the pragmatic considerations involved in lexicalization we discussed in section 4.1 (see also Note 8).

In sum, we have argued that the *push-open* type of V-A form and its separate counterpart constitute a particle verb construction. To put this differently, we have argued that certain types of resultative constructions should be treated as separate forms of particle verb constructions. This view receives a straightforward support from data taken from North Germanic languages. Of particular significance is the fact that unlike in English, the position of a particle can be fixed in these languages. In Swedish, for example, particles are fixed in the pre-object position and do not appear in the post-object position, as the following example shows:

(41)   a. Han kastade ut böckerna.

    he threw out books.the   "He threw out the books."

    b.*Han kastade böckerna ut

    he threw books.the out   (Toivonen 2003:105)

Another fact that has a special significance here is that in Swedish resultative constructions, some adjectives always precede objects, as exemplified in (42) below, and other adjectives always appear in the post-object position, as shown in (43) below. Toivonen (2003: 112-132) claims that this distributional bifurcation in constructions expressing resultant states can be accounted

for if we view the adjective in (42) as a particle, or a "non-projecting word." Some adjectives expressing results are in fact particles, while other adjectives, including the adjective in (43), are not.[10]

(42)   a. Han <u>slog</u> <u>ihjäl</u>   en karl.

       he   beat to.death a   man   "He beat a man to death."

     b. *Han <u>slog</u> en karl <u>ihjäl</u>.

       he   beat a   man to.death   (Toivonen 2003: 22)

(43)   a … och <u>klöst</u>   honom <u>blodig</u>.

       and scratched him   bloody   "… and scratched him bloody."

     b.*… och <u>klöst</u>   <u>blodig</u> honom.

       and scratched bloody him   (Toivonen 2003: 31)

To return to English, the resultative exemplified in (42) corresponds to the *push-open* type of V-A form, while the resultative like the one in (43) corresponds to the usual resultative construction. Since English particles crucially differ from Swedish particles in that they can occur not only in the pre-object position but also in the post-object position, the *push-open* type, which is a particle verb construction, can be realized also in a separate form (e.g. *to push the door open*). That is, due to the "projecting-word" nature of English particles (Toivonen 2003: 166-176), an English counterpart of the pattern in (42b) is acceptable. "True" resultative constructions such as the one in (43a), on the other hand, correspond to many of the resultative constructions in English (e.g. *My father <u>painted</u> the fence <u>white</u>. / My mother <u>bleached</u> the shirt <u>white</u>. / John <u>hammered</u> the metal <u>flat</u>*), i.e. the resultative constructions whose resultative predicates are not one of the disconnection adjectives. Non-particle adjectives cannot occur in the pre-object position unless phrasal lexicalization takes place, as in the case of the *bleach-white* type of V-A form (e.g. **to paint white the fence, *to hammer flat the metal*; see also (43b)).

    Finally, let us note that views similar to our arguments presented in this paper are found in the literature on North Germanic languages. They can be summarized as follows:

(44)   a. The [V-A] adjacent form is a particle verb construction, while the [V…A] separate

       form is a resultative construction.   (Toivonen 2003)

     b. The V-Particle adjacent form cannot be viewed as a morphological compound in

       light of the Righthand Head Rule (see (16)).   (Svenonius 1996, Ramchand

       2008)

     c. "… so-called 'lexical' properties of verbs cannot be confined to a lexical module."

                       (Ramchand 2008: 134)

---

[10] Some adjectives appear in both pre-object and post-object positions, but they have different interpretations in the two positions, as the following examples show (Toivonen 2003: 116-117):

(i) a. … och <u>rycker</u> <u>lös</u> meningar ur   sina sammanhang.

    and pulls free sentences out.of their contexts

    "and pulls free sentences out of their contexts."

  b. Det   <u>springer</u> en vargliknande hund <u>lös</u> på Stocksundsbron.

    there runs   a wolf.like   dog free on S.bridge

    "A dog who looks like a wolf is running free on the Stocksund bridge."

In the pre-object position in (ia), the adjective *lös* is a particle and expresses a resultative meaning, while in the post-object position in (ib), it has a depictive reading. In the latter case, the adjective cannot be regarded as a particle.

The analysis in (44a) has been discussed just above. The view in (44b) is parallel to our claim that though English V-A forms have the property of lexical integrity, they cannot be treated as compound verbs in view of their left-headedness. According to Ramchand (2008: 133), particles in Swedish do occur in a [Particle-Verb] form, as shown in (45b) below. She argues that this is a morphological compound conforming to the Righthand Head Rule, so the Verb-Particle form in (45a) must be formed in a different, non-morphological way. This view is suggested in (44c); the problem of the Righthand Head Rule, which applies to the morphology, would be avoidable if the lexical status of a Verb-Particle form comes from a non-lexical module. A strong piece of evidence for this view is provided by the Norwegian data given in (46) below. Notice the well-formedness of the sentence in (46a) in contrast to the ill-formedness of the one in (46b), which means that the Verb-Particle adjacent form allows adverbial modification of the particle, a property absent from morphological combinations.

(45) a. Det blev   <u>hugget</u>   <u>ned</u>   många träd.

  it   became chopped down many   trees   "Many trees got chopped down."

  b. Det blev   många träd <u>nedhuggna</u>.

  it   became many   trees down.chopped   "Many trees got chopped down."

  (Ramchand 2008: 133)

(46) a. Kari <u>sparka</u> heldigvis   <u>ut</u>   hunden.

  Kari kicked fortunately out the.dog   "Kari fortunately kicked the dog out."

  b. *Kari <u>sparka</u> <u>ut</u> heldigvis   hunden.

  Kari kicked out fortunately the.dog   (Ramchand 2008: 133-134)

5. Conclusions

In this paper, we have taken a close look at English resultative constructions that have alternative forms in which the V and A occur adjacently. On the basis of the morphological, syntactic, semantic, and pragmatic properties of the V-A forms, we have argued that the lexical unit of V-A form is not formed by compounding but arises as an epiphenomenon of the morphology-syntax interaction. The *bleach white* type of V-A form is a lexicalization from a resultative construction, while the *push open* type of V-A form is a type of particle verb construction. We crucially differ from previous studies in our conviction that V-A forms are not a unitary category and that they consist of lexical units formed non-morphologically.

These conclusions confirm the traditional view that genuine verbal compounding is impossible in English, and what appear to be compound verbs are derivatives from various sources. In fact, we have revealed previously unknown ways to form "compound verbs" in English. Traditionally, "compound verbs" in English have been dealt with by means of various word-formation processes such as conversion and back-formation. As we saw in (15), N-V and A-V verbs (e.g. *to stage-manage, to shortcut*) are back-formations or conversions from compound nouns or adjectives, while P-V verbs (e.g. *to download*) are inversions from particle verbs. However, this paper has shown that the means to form "compound verbs" in English may lie outside the morphological component as well as inside it. V-A verbs are either particle verbs or lexicalizations from resultative constructions. The remaining question is: why is it impossible to

form "compound verbs" by compounding per se in English? We leave this question for future research.[11]

Our contribution to theoretical concerns is that we have provided another piece of evidence for the view that lexical units are not equal to morphological constructs. The lexicon can be expanded by non-morphological operations. There exist "words" that are formed outside the morphological component. As we saw in (44c), Ramchand (2008: 134) expresses this view in her claim that "… so-called 'lexical' properties of verbs cannot be confined to a lexical module." Moreover, the existence of non-morphological lexical units like V-A forms strongly speaks for the modification of the notion of Lexical Integrity as a property of X^0 terminals (syntactic atoms) rather than of lexemes (Spencer 2005: 80-81). Such a position is advanced by Ackerman and LeSourd (1997: 99), who claim that "lexical integrity does not hold of lexical items as such, but rather is a property of the zero-level categories specified in lexical representations." Particle verb constructions constitute a classic example of this type of morphology-syntax interaction in English (e.g. Jackendoff 2002), but this paper has revealed the possibility that the notion of lexicalization could be discussed from a similar point of view. That is, lexicalization is not a peripheral phenomenon of merely descriptive value but deserves a serious investigation as an active linguistic mechanism involved in the expansion of the lexicon and in the interaction between morphology and syntax. We may regard Shimamura's studies on phrasal lexicalization (e.g. Shimamura 1986, 2000, 2003) as a starting point for research in such a direction, but a number of significant questions remain to be investigated, including the question of structural and semantic conditions for (phrasal) lexicalization and the distinction between synchronic and diachronic lexicalization. We will tackle on these issues in future works as a necessary step to develop a more articulated theory of lexicalization.

References

Ackema, P. and A. Neeleman 2004, *Beyond Morphology: Interface Conditions on Word Formation*, Oxford University Press, Oxford.

Ackerman, F. and P. LeSourd 1997, "Toward a lexical representation of phrasal predicates," *Complex Predicates*, ed. by A. Alsina, J. Bresnan, and P. Sells, CSLI, Stanford, pp. 67-106.

Adams, V. 2001, *Complex Words in English*, Person Education, Harlow.

Baker, M. C. 1985, "Syntactic affixation and English gerunds," *Proceedings of the 4th West Coast Conference on Formal Linguistics*, pp. 1-11.

Bauer, L. 1983, *English Word-Formation*, Cambridge University Press, Cambridge.

Berg, T. 1998, "The (in)compatibility of morpheme orders and lexical categories and its historical implications," *English Language and Linguistics* 2, pp. 245-262.

Bolinger, D. 1971, *The Phrasal Verb in English*, Harvard University Press, Cambridge, MA.

Booij, G. 2002, "Separable complex verbs in Dutch: A case of periphrastic word formation," *Verb-Particle Explorations*, ed. by N. Dehé, R. Jackendoff, A. McIntyre, and S. Urban, Mouton de Gruyter, Berlin, pp. 21-41.

Booij, G. 2005, *The Grammar of Words: An Introduction to Linguistic Morphology*, Oxford University Press, Oxford.

---

[11] It is interesting to note that Japanese compound verbs show the opposite possibility; Japanese has "lexical compound verbs" and "syntactic compound verbs," which are formed by the process of compounding applied in the lexical and syntactic components respectively (see Kageyama 2009).

Booij, G. 2009, "Compounding and construction morphology," *The Oxford Handbook of Compounding*, ed. by R. Lieber and P. Štekauer, Oxford University Press, Oxford, pp. 201-216.

Botha, R. 1984, *Morphological Mechanisms: Lexicalist Analyses of Synthetic Compounding*, Pergamon Press, Oxford.

Bresnan, J. and S. Mchombo 1995, "The lexical integrity principle: Evidence from Bantu," *Natural Language and Linguistic Theory* 13, pp. 181-254.

Brinton, L. J. and E. C. Traugott 2005, *Lexicalization and Language Change*, Cambridge University Press, Cambridge.

Carstairs-McCarthy, A. 2002, *An Introduction to English Morphology*, Edinburgh University Press, Edinburgh.

Dehé, N., R. Jackendoff, A. McIntyre, and S. Urban (eds.) 2002, *Verb-Particle Explorations*, Mouton de Gruyter, Berlin.

Dowty, D. 1979, *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*, Reidel, Dordrecht.

Fischer, O. 2007, *Morphosyntactic Change: Functional and Formal Perspectives,* Oxford University Press, Oxford.

Hohenhaus, P. 2005, "Lexicalization and institutionalization," *Handbook of Word Formation*, ed. by P. Štekauer and R. Lieber, Springer, Dordrecht, pp. 353-373.

Hoshi, H. 1992, "Circumstantial predicates, PRO, and D-structure adjunction," *English Linguistics* 9, pp. 1-20.

Jackendoff, R. 1990, *Semantic Structures*, MIT Press, Cambridge, MA.

Jackendoff, R. 2002, "English particle constructions, the lexicon, and the autonomy of syntax," *Verb-Particle Explorations*, ed. by N. Dehé, R. Jackendoff, A. McIntyre, and S. Urban, Mouton de Gruyter, Berlin, pp. 67-94.

Kageyama, T. 2009, "Isolate: Japanese," *The Oxford Handbook of Compounding*, ed. by R. Lieber and P. Štekauer, Oxford University Press, Oxford, pp. 512-526.

Kanemoto, M. 2002 "'Kooi-kekka' fukugoogo□kekka koobun tono taisyoo kenkyuu ('Act-result' compounds: A comparative study with resultative constructions)," *Journal of Language and Information Sciences* 7, Association for Language and Information Sciences, The University of Tokyo, pp. 81-98.

Kato, K. and T. Kageyama 1998, "Phrasal compounds and lexical integrity," *English Linguistics* 15, pp. 309-315.

Levin, B. and M. Rappaport, H. 1995, *Unaccusativity: At the Syntax-Lexical Semantics Interface*, MIT Press, Cambridge, MA.

Lieber, R. 2005, "English word-formation processes: Observations, issues, and thoughts on future research," *Handbook of Word Formation*, ed. by P. Štekauer and R. Lieber, Springer, Dordrecht, pp. 375-427.

Lieber, R. 2009, "IE, Germanic: English," *The Oxford Handbook of Compounding*, ed. by R. Lieber and P. Štekauer, Oxford University Press, Oxford, pp. 357-369.

Lipka, L., "Lexicalization and institutionalization in English and German," *Linguistica Pragensia* 1, 1992, pp. 1-13.

Marchand, H. 1969, *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*, Second edition, C.H.Beck, Munich.

McNulty, E. 1988, *The Syntax of Adjunct Predicates*, Doctoral dissertation, University of Connecticut.

Nagano, A. 2009, "Compound verbs in English: Their subclasses and inflectional properties," *JELS* 26 (Proceedings of the 26th Conference of the English Linguistic Society of Japan), pp. 180-198.

Plag, I. 2003, *Word-Formation in English*, Cambridge University Press, Cambridge.

Ramchand, G. C. 2008, *Verb Meaning and the Lexicon: A First-Phase Syntax*, Cambridge University Press, Cambridge.

Sauer, H. 2004, "Lexicalization and demotivation," *Morphology: An International Handbook on Inflection and Word-Formation Vol. 2*, ed. by G. Booij, C. Lehmann, J. Mugdan, and S. Skopeteas, Walter de Gruyter, Berlin, pp. 1625-1636.

Scalise, S. 2008, "Searching for universals in compounding," paper presented at the symposium "Morphology and its neighboring areas" in the 136th Conference of the Linguistic Society of Japan.

Shimamura, R. 1986, "Lexicalization of syntactic phrases," *English Linguistics* 3, pp. 20-37.

Shimamura, R. 2000, "Lexicalization of syntactic phrases: The case of genitive compounds like *woman's magazine*," [http://coe-sun.kuis.ac.jp/public/paper/outside/shimamura2.pdf].

Shimamura, R. 2003, "On lexicalized phrases," *Empirical and Theoretical Investigations into Language: A Festschrift for Masaru Kajita*, ed. by S. Chiba et al., Kaitakusha, Tokyo, pp. 631-646.

Spencer, A. 2005, "Word-formation and syntax," *Handbook of Word Formation*, ed. by P. Štekauer and R. Lieber, Springer, Dordrecht, pp. 73-97.

Svenoius, P. 2008, "The verb-particle alternation in the Scandinavian languages," ms., downloaded from [http://ling.auf.net/lingBuzz/000046].

Taniwaki, Y. 2006, "Resultative verb-adjective combinations as lexical compounds," *Lexicon Forum* 2, ed. by T. Kageyama, Hituzi Shobo, Tokyo, pp. 251-280.

Toivonen, I. 2003, *Non-Projecting Words: A Case Study of Swedish Particles*, Kluwer, Dordrecht.

Washio, R. 1997, "Resultatives, compositionality and language variation," *Journal of East Asian Linguistics* 6, pp. 1-49.

Williams, E. 1981, "On the notions 'lexically related' and 'head of a word'," *Linguistic Inquiry* 12, pp. 245-274.

Zeller, J. 2002, "Particle verbs as heads and phrases," *Verb-Particle Explorations*, ed. by N. Dehé, R. Jackendoff, A. McIntyre, and S. Urban, Mouton de Gruyter, Berlin, pp. 233-267.

Are there coordinate compounds?[*]

Susanna Padrosa-Trias
Centre de Lingüística Teòrica, UAB
susannapt@gmail.com

1. Introduction

A number of studies dedicated to compounding acknowledge the existence of the so-called coordinate compounds, which can be roughly defined as compounds in which there is a relation of coordination between the two constituents (e.g. Bauer 2001, 2008, Bisetto & Scalise 2005, Olsen 2001, 2004). Coordinate compounds are usually divided into two subtypes, which are exemplified in (1) and (2) for English.[1]

(1)   actor-director

        player-coach

        jazz-rock

(2)   mother-child (relationship)

        doctor-patient (gap)

        mind-body (problem)

The basic claim of the paper is the denial of such a type of compounding. It is argued that what are generally called coordinate compounds are cases of asyndetic syntactic coordination. It is shown that coordinate structures can, nonetheless, be interpreted as compounds under special circumstances.

The paper is structured as follows. Section 2 includes a review of two different approaches to coordinate compounding: Bisetto & Scalise's (2005) and Olsen's (2001, 2004). Section 3 contains my proposal - according to which nominal coordinate compounds of the NN type are nonexistent - and section 4 extends the proposal to verbal categories. Finally, section 5 summarizes the main findings of this study.

---

[1] Note that *coordinate compounds* is only one of the labels used in the literature to make reference to the forms in (1) and (2) jointly. Other labels are *copulative compounds* (cf. Olsen 2000, 2001, 2004), *appositional compounds* (Bauer 2001, 2008) and *appositive compounds* (Booij 2005). Sometimes distinct labels are used for the compounds in (1) and those in (2): *appositional and coordinative compounds*, respectively (Plag 2003), *appositional and relational compounds*, respectively (Wälchli 2005). In this paper the label *coordinate compounds* has been chosen to refer to the forms in (1) and (2) jointly, except where Olsen's understanding of such forms is described.
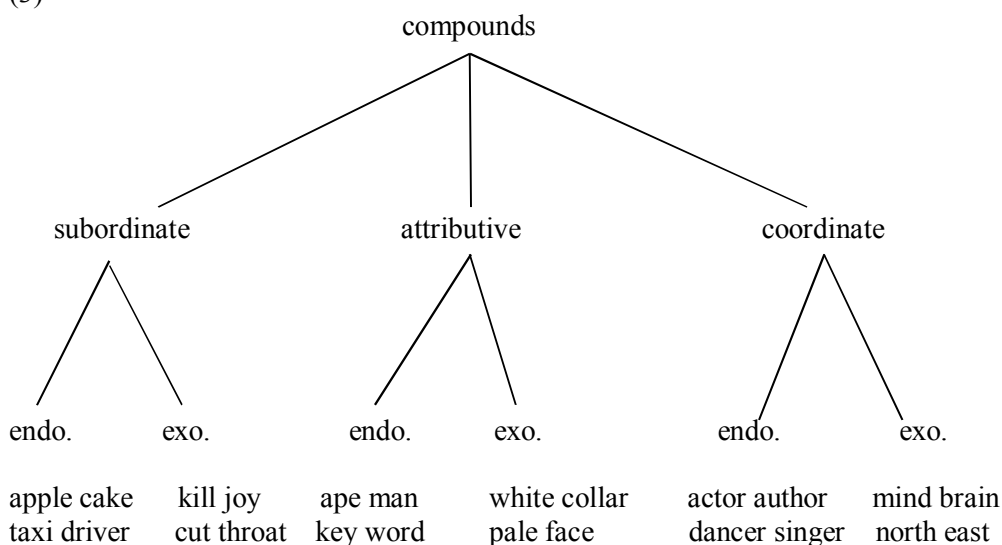
2. Coordinate compounding

This section briefly presents how the so-called coordinate compounds are understood in Bisetto & Scalise (2005) and in Olsen (2001, 2004). This will allow us to gain a general idea of how coordinate compounds are treated in the literature.

2.1. Bisetto and Scalise (2005)

Bisetto & Scalise (2005)[2] provide a classificatory scheme for compounding, which is as follows:

(3)

```
                          compounds
              ┌───────────────┼───────────────┐
         subordinate      attributive       coordinate
          ┌────┴────┐      ┌────┴────┐      ┌────┴────┐
        endo.      exo.  endo.      exo.  endo.      exo.

      apple cake  kill joy  ape man  white collar  actor author  mind brain
      taxi driver cut throat key word pale face     dancer singer north east
```

The first level of analysis takes into account the grammatical relation that holds between the constituents of the compound. By this criterion, three macro-types of compounds are identified, each defined by a different relation. One of them is a relation of subordination, which can be found in compounds like *car-driver*, where *car* is understood as the internal argument of *drive*, *book cover*, interpreted as the 'cover of a book', and *catfood*, understood as 'food for cats'. This type of relation gives rise to subordinate compounds, which are contrasted with the two other macro-types: attributive compounds and coordinate compounds. Attributive compounds are characterized by a relation of attribution: the first element expresses a property which is

---

[2] Bisetto & Scalise's (2005) classification is revised in Scalise & Bisetto (2009). In a nutshell, Scalise & Bisetto (2009: 50) add a further level of analysis into their previous classification of compounds "in order to account for the different semantic/interpretative relations that come into place between the constituents of the compounds in each class". Subordinate compounds and attributive compounds, but not coordinate compounds, are further subdivided. For example, subordinate compounds are divided into ground and verbal-nexus compounds. Ground compounds correspond to what in the literature on compounding has generally been referred to as primary/root compounds (e.g. *windmill*), whereas verbal-nexus compounds include those compounds in which the head is deverbal and the non-head can be either an argument (often called synthetic compounds, e.g. *bookseller*) or an adjunct (e.g. *street seller*). Since there are no changes regarding coordinate compounds, the discussion to follow is based on Bisetto & Scalise (2005).

attributed to the second element, as in *blue cheese* and *pale face*. Coordinate compounds are characterized by a coordinating relation, and are defined as follows:

(4)   "Coordinate compounds are those formations whose constituents are tied by the conjunction «and». (…) From a semantic point of view, such compounds can be considered as having two heads (*poet painter* is both a «poet» and a «painter»)."

Bisetto & Scalise (2005: 327)[3]

The three compounding macro-types[4] are then defined by a second criterion, which is characterized by the presence or absence of a head. This second criterion divides each macro-type into two sub-types: endocentric and exocentric.

Scalise & Guevara (2006) observe that presence vs. absence of a head constituent can be ambiguous between a formal head and a semantic head, which they define as follows (p. 190):

(5)   "The formal head of a compound is the constituent which shares with –and percolates to- the whole compound all of its formal features: lexical category and subcategorization frame. The whole compound, thus, is expected to have the same distributional properties of its formal head."

"The semantic head of a compound is the constituent which shares with –and percolates to- the whole compound all of its lexical-conceptual information (LCS in short, following Jackendoff 1990 and Lieber 2004). The whole compound, thus, is expected to be a hyponym of its semantic head."

Scalise & Guevara (2006) claim that endocentricity obtains in those compounds where the formal head and semantic head coincide, as in *capostazione* (lit. master+station, 'station master') in which the semantic head (a *capo*, which is a hyperonym of a *capostazione*) is the same as the formal head (the masculine gender of the compound comes from *capo*: $[[capo]_{masc}[stazione]_{fem}]_{masc}$). When the two heads do not coincide, then the compound is exocentric. In their terms (p. 192):

(6)   "An *endocentric compound* has at least one formal head and at least one semantic head. If a compound has only one formal head and only one semantic head, then the two must coincide.

If a compound realises any of the remaining possibilities, it will be considered to be *exocentric*."

---

[3] Scalise & Bisetto (2009: 46) give a similar definition, which I quote because it will become relevant in section 3.1.: "*From a semantic point of view, these compounds can be considered to be characterized by two heads (painter-poet is both a 'poet' and a 'painter') even though, **as claimed by Bloomfield (1933), only one of the nouns can act as the head. As a general rule, only one of the nouns can be pluralized, and, in those languages where gender is relevant, it is precisely that noun that confers the gender on the compound formation**.*" [bold: SPT]

[4] The tripartite classification of compounds is allegedly reinforced in Scalise, Bisetto & Guevara (2005), in which it is argued that each macro-type of compounding has a different selection mechanism. That is, the head of the compound is supposed to select the non-head differently in each of the three macro-types. See Padrosa-Trias (forthcoming) for some criticisms of this proposal.

This notion of head has consequences for the understanding of coordinate compounds. Initially, coordinate compounds, despite inflection being placed on both elements of the compound in languages like Italian, were taken to have one head, which was determined by the canonical head position of the language in question. For example, English would have the head on the right, e.g. *actor-director*, and Italian would have the head on the left, e.g. *bar pasticceria* 'bar-pastry shop'. This view is endorsed in Scalise, Bisetto & Guevara (2005), which is later changed to incorporate the revised notion of head and endocentricity/exocentricity in (5) and (6), according to which coordinate compounds have two heads (Scalise & Guevara 2006: 191). In short, the uneasiness about the notion of head and the change in coordinate compounds from having one head to two heads suggests that the structure of such compounds is not crystal clear.
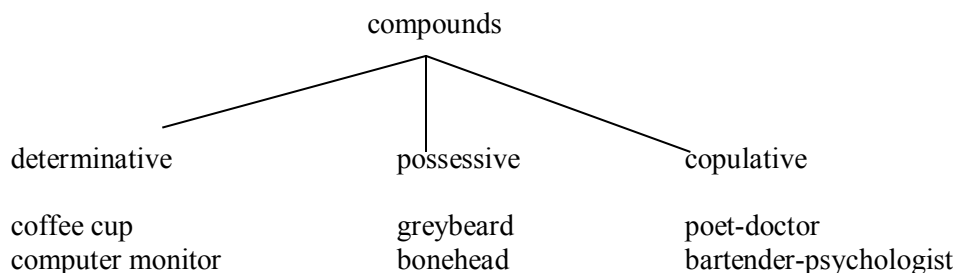
After having presented Bisetto & Scalise's (2005) compounding scheme together with some refinements to their classification, one is in a position to say that Bisetto and Scalise would treat the coordinate compounds in (1) as headed (with two heads: endocentric) while those in (2) would be treated as headless (exocentric).

Let us now turn to Olsen's (2001, 2004) classification of compounds and to her understanding of coordinate compounds.

2.2. Olsen (2000, 2001, 2004)

Olsen (2000, 2001, 2004) follows the classification of compounds used by the early grammarians of Sanskrit, according to which compounds are divided into three major types:

(7)

```
                         compounds
             _____/    |    _____
            /                |                \
    determinative        possessive         copulative

    coffee cup           greybeard          poet-doctor
    computer monitor     bonehead           bartender-psychologist
```

Although the three major types of compounds in (7) do not exactly correspond to Bisetto & Scalise's (cf. 3) three macro-types, the determinative, possessive and copulative compounds of (7) can be subsumed into the subordinate, attributive and coordinate compounds of (3) respectively. Let us consider each type in turn, placing special emphasis on copulative compounds. Determinative compounds are those in which the first element restricts the denotation of the second element, the head: e.g. a *coffee cup* is a type of cup, one for coffee. Possessive compounds also display a modifier-head relation which, in this case, denotes a property which is attributed to an external entity: a *greybeard* is, for example, a seal which has a grey beard. Copulative compounds are defined as follows:

(8) "Copulative compounds (or *pseudo-dvandvas* (…)) are compounds in which the individual constituents are equally predicated of the entity to which the compound as a whole refers (…). Some recent coinages are *actor-houseguest*, *gangster-businessmen*, *host-mediator*, *explorer-anthropologist*, *tent-office* and *Kosher-Cajun*. An *actor-houseguest* is someone who is both an actor and a houseguest; a *tent-office* is something that is both a tent and an office and *Kosher-Cajun* refers to a type of cuisine possessing the characteristic attributes of both manners of preparing food."

<div align="center">Olsen (2000: 908)</div>

"Copulative compounds encompass a coordinative relationship between the two constituents such that both concepts are attributed simultaneously to one individual: 'poet-doctor' is someone who is both a 'poet' and a 'doctor'."

"(…) the basic copulative pattern carries the meaning 'an x that is simultaneously A and B' (…)."

<div align="center">Olsen (2001: 279, 297)</div>

As can be seen, the subordinate, attributive and coordinate compounds of (3) include the determinative, possessive and copulative compounds of (7) respectively, but also other types of compounds. For example, some attributive compounds refer to an entity which is characterized by the property expressed by the compound (e.g. *pale face*), in the same way as possessive compounds, but there are also other attributive compounds in which the head is modified by the non-head, with no reference to a third entity (e.g. *ape man*), unlike possessive compounds. Further differences between the two classifications will not be pursued, since my main concern is the characterization of the so-called coordinate compounds, or copulative compounds in Olsen's terms.

Olsen (2001, 2004) argues that the three compound patterns displayed in (7) are subsumed into the same compound template in languages like English and German:

(9) $[Y + X]_x$ (Y and X being open lexical categories)

This formal scheme is implemented semantically in Olsen (2004: 89f), in which it is stated that the two predicates that constitute the compound stand in an underspecified relation to one another. The relation is taken as a variable whose content is predicted by the meaning of the compounding elements or inferred from a contextually relevant aspect. Let us consider how the determinative, possessive and copulative readings are derived given a single compound template. Possessives are assimilated to determinatives with the only difference that exocentric possessives involve a process of meaning extension. Copulative compounds are a semantic subset of the template displaying the 'and' relation between the constituents of the compound. The differences between the three compound types then have to do with the interpretational option chosen. To be more precise, the determinative/possessive reading is obtained when the underspecified relation is instantiated as a modifier-head relation, as in *coffee cup*, and the copulative reading is obtained when the relation is instantiated as an identity relation, as in *actor-houseguest*. There are some compounds in which both relations can be instantiated with the result that the same compound can be interpreted both as a determinative/

possessive compound and as a copulative compound. As Olsen (2004: 91) notes, *bartender-psychologist* is a case in point. It can be understood as a 'psychologist for a bartender', a 'psychologist that treats a bartender', a 'psychologist who looks like a bartender', among other determinative readings, and also as a person who is both a psychologist and a bartender, namely the copulative reading.

By proposing the $[Y + X]_x$ template, Olsen predicts that regular compounding in Germanic will be right-headed formally. The prediction seems to be corroborated. Copulative compounds in English have the plural inflection on the second constituent:

(10)   a. (…) *the writer-directors* (John Musker and Ron Clements…)

     b. (Disney's) *attorney-archivists*

The plural inflection is also placed on the second constituent in German (11a) and the second constituent determines the gender of the whole compound (11b):

(11)   a. *die Linguist-Psychologen, die Ingenieur-Studenten*

     b. *der Baby-Bastard, der Opfer-Zeuge*

In Olsen's view, copulative compounds are hierarchically structured with binary branching, namely they do not have a flat structure. For instance, *songwriter-producer-arranger-friend* is given the structure below:

(12)   [[[[songwriter] producer] arranger] friend][5]

Olsen notes that copulative compounds can occur on their own and in an embedded position. Regarding unembedded copulatives, some semantic patterns are more common than others. Some compounds refer to things, as in *tent-office* and *comedy-drama*, but the most productive semantic pattern denotes people by naming their professions, as in *writer-director*, *singer-guitarist*, and *editor-publisher*. Following Olsen (2001: 305, 2004: 88), the two compounding elements together form a complex concept that is added to one's ontological system of objects. If the two elements in a copulative compound cannot create a concept referring to a coherent entity in one's ontological system of individuals, then the result is ungrammaticality, as in (13).

(13)   *The artist-instrument thrived on irony.[6]

Concerning embedded copulatives, they can be inserted in structures in which the head licenses a semantically coordinate complex argument. In some cases, the head allows a complex argument which displays a 'between' relation between its members: *predator-prey battles* are battles between predators and preys. In other cases, the head can be a collective term whose constitution is specified by the elements of the embedded copulative: *a man-wife team* is a team made up by a man and his wife (for other contexts in which embedded copulatives are allowed, see Olsen 2001: 298-301). In short, regarding the examples just mentioned, a copulative compound is embedded into a

---

[5] The examples in (10), (11), and (12) are taken from Olsen (2001: 293).

[6] This example (borrowed from Olsen 2004: 88) is contrasted with syntactic coordination, in which the same predicates are used but now they are predicated of an individual.

     (i)   Warhol, the pop artist and (the) instrument of the masses, thrived on irony.

                                                                    Olsen (2004: 88)

determinative compound: e.g. *[[predator-prey] battles]*, and the embedded copulative (*predator-prey*) is licensed by the semantic requirements of the head (*battles*).

Turning to the question of how to treat the compounds in (1) and (2), Olsen would call them copulative and would distinguish them by their embedded (2) vs. unembedded (1) nature.

After having presented two views on coordinate/copulative compounds (Bisetto & Scalise's and Olsen's), my proposal concerning their status follows.

3. The proposal

This section presents my proposal regarding the nature of the alleged coordinate/copulative compounds illustrated in (1) and (2). The forms in (1) are dealt with in section 3.1. while the forms in (2) are discussed in section 3.2.

3.1. Endocentric coordinate compounds / unembedded copulative compounds

I depart from Bisetto & Scalise's (2005) and Olsen's (2000, 2001, 2004) conception of coordinate/copulative compounds substantially. They all understand that the compounds in (1) refer to an *entity* which is both A and B, A and B being the two members of the compound: e.g. an *actor-director* is *somebody* who is both an actor and a director (see Olsen's definitions in (8)). If all forms in (1) refer to an entity outside the coordinate structure, they cannot be endocentric, as Bisetto & Scalise claim, but exocentric (see Levi 1978 for a similar view).[7] If this reasoning is correct, there are no endocentric coordinate compounds, and the coordinate compounds in (1) should be labelled exocentric like those in (2).

However, I want to argue that there are no exocentric coordinate compounds either. In my view, a true coordinate relation (i.e. an entity having properties of both A and B) can only be established in syntax, not in morphology where compounding takes place (see Haspelmath 2004 for a broad view on coordinating constructions).[8] Support for the proposal that coordination is syntactic (as opposed to morphological) comes from authors like Bresnan & Mchombo (1995) who discuss some tests which show that the internal structure of words behaves differently from that of phrases. One of the tests is conjoinability: syntactic objects can be conjoined by the coordinator *and* while morphological objects cannot.[9] Accordingly, NN forms with a coordinate relation cannot be treated as compounds. My proposal is that they are cases of asyndetic

---

[7] Levi (1978: 93-94) believes that, despite the compounding nouns being in a coordinate relation, the resulting compound (or the 'complex nominal' in her terms) is exocentric because neither noun is the head semantically. She proposes an underlying relative clause whose head is deleted. For example, *speaker-listener* is derived from 'person who is (both) a speaker and a listener', with *person* being deleted.

[8] I believe that (alongside a generative syntactic component) there is a generative morphological component responsible for word formation processes like compounding.

[9] Some apparent counterexamples seem to involve ellipsis/deletion and cannot then be treated as conjunctions of parts of words (but see Ackerman & LeSourd 1997 and Lieber & Scalise 2006 for a different view).

syntactic coordination, with an implicit conjunction between the two nouns. The forms in (1) cannot then be coordinate compounds of any type. The same conclusion is reached by Adams (2001: 82), who does not consider similar forms compounds on the grounds that expressions with 'coordinated elements are phrases' (in this case the coordinator would be implicit).

The forms in (1) can, nonetheless, be treated as compounds, as endocentric single-headed compounds, determinative compounds in Olsen's terms: the second noun is the head formally (e.g. plural marker is attached to it) and semantically (it is a hyperonym of the compound as a whole); and the first noun restricts/modifies the head. As a result, the compound instantiates a modification/subordination relation: the compound denotes a subset of the set of entities denoted by the head noun, which is given some properties by the first noun. As defined by native speakers, a *player coach* is 'a coach who is also a player on the team', 'a coach that plays with the team' and *jazz rock* is 'rock with some characteristics of jazz'.

The facts observed in (10-12) are easily captured if compounds like those in (1) are determinative, rather than coordinative (contra Olsen). The second constituent bears the plural inflection in the examples in (10) and (11a) and determines the gender of the whole compound in (11b), facts which follow if the second compounding element is the formal head, and which are hard to explain if the compound is coordinate. That is, the formal right-headedness of the compounds is expected if they are determinative but is not expected if they are coordinate (cf. Bloomfield's 1933 remarks in footnote 3). If a coordinate relation were present between the compounding elements, plural inflection would be expected on both elements, contrary to fact. The hierarchical structure depicted in (12) for *songwriter-producer-arranger-friend* and the compound template [Y + X]$_x$ shown in (9), proposed by Olsen, can accommodate determinative compounds better than the alleged coordinative compounds: in the case of determinatives the relationship between the constituents of the compound is subordinative, with a modifier-head relation, and in the case of copulatives it is coordinative, with a symmetrical relation.[10] An asymmetrical relation seems to be instantiated both in (12) and (9), thus favouring the determinative type of compound. In addition, if there were a relation of coordination between the compounding constituents, the two nouns would equally be hyperonyms of the compound. Alleged coordinate compounds like *player coach*, though, are interpreted as determinative compounds by native speakers (see above). In short, the forms in (1) fit the determinative pattern of compounding both formally and semantically, while they prove problematic to conform to an alleged coordinate pattern of compounding.

Notice that my proposal does not deny a sequence of two nouns the possibility of having a coordinate reading. My claim is that when such a reading is present, one is dealing with a syntactic construction with asyndetic coordination (and not with compounding). In other words, an NN sequence can be interpreted as encoding a relation of coordination, in which case it is a syntactic construction, or as encoding a modifier-

---

[10] For a definition of coordination in terms of (a)symmetry, see Haspelmath (2004: 35f).

head relation, in which case it is a determinative compound. My proposal is summarized in the following scheme:

(14)   NN → coordinate reading → not a compound

NN → modifier-modified reading → a compound

Some NN sequences can only be interpreted as determinative compounds. This is the case of the compounds whose first element denotes the gender of the noun in second position: *maid-servant*, and *she-goat*. It seems that a coordinate relation is possible when the two coordinated elements can equally contribute new information to the construction by their being semantically parallel. These requirements are not satisfied by compounds whose first element is a gender marker (*she-goat*), but seem to be satisfied by forms denoting two job titles (e.g. *actor-director*) or two types of devices/machines (*washer-dryer*) although not always (e.g. *fighter-bomber*), according to native speakers' judgments. What these results suggest is that two apparently coordinated nouns can be interpreted as coordinate but also as a modifier-modified structure, the final interpretation probably being subject to the speaker's knowledge of the world.

Before delving into the forms in (2) in the next section, let us consider how the proposal put forth in this section can deal with the ungrammaticality of the sentence in (13). According to my proposal, its ungrammaticality is not due to the fact that the complex concept created by *artist* and *instrument* does not refer to a coherent individual in one's ontological system of individuals, as Olsen claims, but to the inability of *artist-instrument* to conform to a determinative compound in the sentence in (13). That is, in my understanding, *artist-instrument* cannot encode a coordinate relation between its constituents if it is to be understood as a compound (recall that coordination is a sign of syntax and not of morphology, i.e. compounding), but it can be a determinative compound with a modifier-head relation. The interpretation of an instrument that serves as an artist or that is like an artist in some aspect seems plausible (which would be consistent with the interpretation of a determinative compound). The ungrammaticality of (13), though, follows from the fact that, out of context, *instrument* is understood as a tool (an object) and not as a person, an interpretation that clashes with the semantic requirements of the verb *thrive* and the sentence in general. The semantics of the construction in which *artist-instrument* is placed do not agree with the expectations created by the determinative compound regarding its semantics: *thrive on irony* requires an agentive subject, which clashes with the default reading of *artist-instrument* as an object. A kind of garden path effect seems to cause the ungrammaticality of the sentence.

3.2. Exocentric coordinate compounds / embedded copulative compounds

The forms in (2), repeated below for convenience, are treated as exocentric coordinate compounds in Bisetto & Scalise (2005) and as embedded copulative compounds in Olsen (2001, 2004). According to their view, the two members of the compound characterize an entity outside the compound, with which they stand in a particular relationship, as in *the mind-body problem*, understood as the problem between the mind

and the body. Some authors distinguish different subtypes of such compounds. For instance, Bauer (2008) distinguishes translative compounds (*the Wellington-Auckland flight*) from co-participant compounds (*parent-child relationship*). In the former the order of the elements makes a difference in meaning since there is a starting point and a finishing point, and in the latter there is some interaction among the participants. However, the position taken by the aforementioned authors cannot be maintained if coordinate compounds do not exist, as has been claimed in the previous section. Let us now consider how the forms in (2) can be analysed in agreement with the proposal according to which there are no coordinate compounds.

(2)   mother-child (relationship)

        doctor-patient (gap)

        mind-body (problem)

Although an NN sequence with a coordinate relation is a phrase (as already discussed above in relation to the nature of the objects in (1)) and cannot be a compound by itself, it can be incorporated in the non-head position of a compound, if permitted by the head. This is the case of the compounds in (2), which in my view are endocentric compounds with a subordinate relation between the head and the non-head. My proposal is that the forms in (2) are compounds not by virtue of the coordinate relation established between the elements constituting the phrase (as has generally been assumed) but by virtue of the subordination relation established between the phrase in the non-head position (which acts as a simplex word, cf. Ackema & Neeleman 2004) and the noun in head position. To illustrate the point, in *mind-body problem*, *problem* is the head of the compound and *mind-body* is its non-head, which happens to be a syntactic phrase turned into a word and inserted in the non-head position of the compound. The specific relation between the elements of the compound is determined by the semantics of the head (cf. e.g. Pustejovsky 1995). This type of compound is possible when the head licenses a complex coordinate argument. Recall from section 2.2. that Olsen identifies different types of heads that allow a coordinate phrase in the non-head position of the compound. For example, the collective term *team* allows the phrase *man-wife* to specify the content of the *team* in *a man-wife team*. Some examples in which the head permits a complex coordinate argument are given below:

(15)   the angel-beast division

        father-daughter dance

        the Cadbury-Schweppes business

        Wellington-Auckland flight

        the nature-nurture debate

        love-hate relationship

Some support for my proposal comes from the observation that if the compounds in (2) or (15) were exocentric compounds, as Bisetto & Scalise (2005) claim, they would be quite different from other compounds that are classified as exocentric in their system, such as *butterfingers* and *redhead*. These two compounds are said to be exocentric

because their referent (the 'semantic head') is not determined by *fingers* and *head*,[11] but by an entity outside the compound, i.e. a type of person. However, Bisetto & Scalise's explanation for exocentricity cannot be extended to any of the examples in (2) or (15). For example, *mind-body* does not uniquely refer to a problem (only *mind-body problem* does). In my analysis, *mind-body* just means 'mind and/or/… body' and can be combined within an endocentric compound with any noun to its right: for instance, *mind-body question*, *mind-body relationship*, *mind-body discussion* and *mind-body exhibition*.[12] In contrast, it is impossible to combine a compound like *redhead* with a noun to its right that refers to the semantic head of *redhead* (e.g. *person*), since it would be semantically superfluous (i.e. the word 'person' is already implied): **redhead person*.

In short, what Bisetto & Scalise and Olsen understand for coordinate/copulative compounds can only exist in the non-head position of a subordinate/determinative compound. In this position a coordinate relation (understood as syntactic and not as morphological) can be established thanks to the semantics of the head outside the coordinate relation.

4. Extending the analysis

The analysis proposed for the nominal forms in (1) and (2) can also be applied to verbal and adjectival categories in English, and is intended to be applicable to other languages as well. While verbal VV sequences in English will be briefly presented below, the reader is directed to Padrosa-Trias (2010) for some discussion on AA sequences in English (e.g. *devilish-holy, cruel-compassionate expression*) and for the analysis of parallel examples in Catalan (e.g. *bomber escalador* (firefighter climber) 'a firefighter who can also work as a climber', *vol Àustria-Hongria* (flight Austria-Hungary), *tractat hispano-americà* (treaty Hispano+American)). Let us now consider how VV forms with an apparently coordinate relation between the two constituents are treated following my proposal.

Regarding $[VV]_V$ forms in English, the general consensus is that they are not compounds (Selkirk 1982) or are regarded as exceptional compounds (Spencer 2003) because they are argued to be the result of backformations. They may be related to nominal or adjectival forms: *dive-bomber*$_N$ ~ *dive-bomb*$_V$ and *dry-cleanable*$_A$ ~ *dry-clean*$_V$. However, the grammar is unlikely to result in an acceptable object, namely a $[VV]_V$ compound, if the grammatical principles do not allow such a type of object, which explains why I consider it a compound (cf. Booij 2005, Plag 2003). Some examples follow:

---

[11] They could act as heads in a 'metonymy' analysis, though.

[12] One question that may arise from the previous discussion, though, is why a phrase, without an overt coordinator, is usually odd at best when used syntactically, but fine in the non-head position of a compound (?? *mind-body is an interesting problem*). A tentative answer could be that a syntactic phrase must omit some material if it is to appear in the non-head position of a compound, as has been argued for telegraphic speech in newspaper headlines (see Ackema & Neeleman 2004: 123, fn. 10 for similar discussion), whereas such material must be present in syntax.

(16)   crash-land, dive-bomb, drink-drive, drop-kick, dry-burn, fly-drive, freeze-dry, shrink-wrap, slam-dunk, sleep-walk, stir-fry, and strip-search.

As already noted in the previous section, I understand complex forms with a coordinate relation not as compounds, but as phrases. If the forms in (16) involve a true coordinate relation, they cannot be included in the study of English compounding. Although the presence of asyndetic coordination is a real possibility for some forms (e.g. *stir-fry*), speakers' interpretations show that this is not the only reading available. The forms in (16) can also be analysed as compounds with the second verb being the head formally (e.g. past tense inflection attaches to it) and semantically (i.e. the compound denotes a subtype of the type of action expressed by the second verb), and the first verb being a kind of manner/temporal modifier. The result is that such compounds are endocentric single-headed compounds (of a subordinate/determinative nature). Accordingly, *to dive-bomb* is expected to mean 'to bomb in a diving fashion/when diving', that is, a type of bombing. This expectation agrees with the interpretation given by speakers.

5. Conclusions

In this paper I have argued for the non-existence of coordinate (either endocentric or exocentric)/copulative compounds (either unembedded or embedded) in Bisetto & Scalise's (2005) and Olsen's (2001, 2004) terms. My proposal is that such alleged compounds are cases of asyndetic syntactic coordination. Coordinate structures can, though, be interpreted as compounds if one element is taken as the head and the other as the non-head. This is the case of the examples in (1): for example, a *player coach* is a type of coach. It is also shown that coordinate structures can be inserted in the non-head position of a compound, which is the case of the examples in (2): for example, *the mind-body problem* is a kind of problem, one which has to do with the mind and the body. In short, I have claimed that the forms which are traditionally called coordinate compounds have the following structure: [[non-head] head], with the non-head being filled with a single word, as in *[[player] coach]*, or with a coordinate phrase that acts as a single word, as in *[[mind-body] problem]*.

References

P. Ackema, and A. Neeleman 2004, Beyond Morphology: Interface Conditions on Word Formation, Oxford University Press, Oxford.

F. Ackerman, and P. LeSourd 1997, Toward a Lexical Representation of Phrasal Predicates, A. Alsina, J. Bresnan, and P. Sells (eds.), Complex Predicates, CSLI Publications, Stanford, 67-106.

V. Adams 200l, Complex Words in English, Longman, Harlow.

L. Bauer 2001, Compounding, M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.), Language Typology and Language Universals, volume 1, Walter de Gruyter, Berlin, pp. 695-707.

L. Bauer 2008, Dvandva, Word Structure 1, pp. 1-20.

A. Bisetto, and S. Scalise 2005, Classification of compounds, Lingue e Linguaggio 2, pp. 319-332.

L. Bloomfield 1933, Language, George Allen & Unwin, London.

G. Booij 2005, The Grammar of Words, Oxford University Press, Oxford.

J. Bresnan, and S. Mchombo 1995, The Lexical Integrity Principle: Evidence from Bantu, Natural Language and Linguistic Theory 13, pp. 181-254.

M. Haspelmath 2004, Coordinating Constructions, John Benjamins, Amsterdam.

R. Jackendoff 1990, Semantic Structures, MIT Press, Cambridge, Mass..

J. Levi 1978, The Syntax and Semantics of Complex Nominals, Academic Press, New York.

R. Lieber 2004, Morphology and Lexical Semantics, Cambridge University Press, Cambridge/New York.

R. Lieber, and S. Scalise 2006, The Lexical Integrity Hypothesis in a New Theoretical Universe, Lingue e Linguaggio 1 [Special Issue on Lexical Integrity], 7-32.

S. Olsen 2000, Composition, G. Booij, Ch. Lehmann, and J. Mugdan (eds.), Morphology: An International Handbook on Inflection and Word-Formation, Walter de Gruyter, Berlin, pp. 897-915.

S. Olsen 2001, Copulative compounds: a closer look at the interface between morphology and syntax, G. Booij, and J. van Marle (eds.), Yearbook of Morphology 2000, Kluwer, Dordrecht, pp. 279-320.

S. Olsen 2004, Coordination in morphology and syntax: the case of copulative compounds, A. ter Meulen, and W. Abraham (eds.), The Composition of Meaning, John Benjamins, Amsterdam/Philadelphia, pp. 17-38.

S. Padrosa-Trias 2010, Complex Word-Formation and the Morphology-Syntax Interface, PhD dissertation, UAB,.

I. Plag 2003, Word-Formation in English, Cambridge University Press, Cambridge.

J. Pustejovsky 1995, The Generative Lexicon, MIT Press, Cambridge, Mass.

S. Scalise, and A. Bisetto 2009, The classification of compounds, R. Lieber, and P. Štekauer (eds.), The Oxford Handbook of Compounding, Oxford University Press, Oxford, pp. 34-53.

S. Scalise, and E. Guevara 2006, Exocentric compounding in a typological framework, Lingue e Linguaggio 2, 185-206.

S. Scalise, A. Bisetto, and E. Guevara 2005, Selection in Compounding and Derivation, W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, and F. Rainer (eds.), Morphology and its Demarcation. John Benjamins, Amsterdam/Philadelphia, pp. 133-150.

E. Selkirk 1982, The Syntax of Words, MIT Press, Cambridge.

A. Spencer 2003, Does English Have Productive Compounding?, G. Booij, J. De Cesaris, A. Ralli, and S. Scalise (eds.), Topics in Morphology: Selected Papers from the Third Mediterranean Morphology Meeting, IULA, Barcelona, pp. 329-341.

B. Wälchli 2005, Co-compounds and Natural Coordination, Oxford University Press, Oxford.

A morphological solution to agreement puzzles in Slavic

AKIRA WATANABE

University of Tokyo

akirawat@l.u-tokyo.ac.jp

1. Introduction

In this paper, I propose that underspecification in narrow syntax can be restored to full specification in morphology. Restoration to full specification in morphology gives rise to what one may call disguised syntactic underspecification, in the sense that underspecification is not transparently visible on the PF side.

Syntactic underspecification is motivated by the predictability of feature values. I would like to suggest that in each case, underspecification is keyed to one particular kind of predictors of feature values as a parametric option. On the morphological side, I claim that missing predictable feature values are supplied in morphology only when the predictor of the value is a feature available in morphology. Thus, the type of the value predictor serves as a diagnosis of disguised syntactic underspecification.

The phenomenon to be taken up is failure of agreement with subjects that have a higher numeral in a subset of Slavic languages including Czech, Polish, Serbo-Croatian, and Slovene (see Veselovská [1] for Czech, Rappaport [2] and Rutkowski and Maliszewska [3] for Polish, Franks [4] and Wechsler and Zlatić [5] for Serbo-Croatian, and Marušič and Nevins [6] for Slovene; see also Corbett [7] and Franks [8]). A Czech example is given in (1) to illustrate the point.

(1) Těch   pět hezkých   dívek jelo.   Czech
     this.gen.pl five beautiful.gen.pl girl.gen.pl traveled.n.sg (Franks [8], p. 137)

Here, the neuter singular form of the verb is obligatory, despite the fact that the head noun takes the (genitive) plural form. It is a mistake to attribute the lack of plural agreement directly to the genitive marking of the head noun, since Russian allows plural agreement despite genitive marking, as in (2).[1]

(2) Pjat' krasivyx devušek   prišli.   Russian
     five   beautiful girl.gen.pl arrived.pl   (Franks [8], p. 106)

---

[1] Neuter singular is also possible in Russian, analysis of which is beyond the scope of this paper. See Bošković [9], Franks [4, 8], Pereltsvaig [10], and the references cited there for various possibilities. See also Corbett's [7] p. 215 quantitative corpus data, which exhibits a huge difference between Russian and the set of Slavic languages that concern us here.

Let me also note that care must be taken to exclude from consideration subjects made exceedingly long by modifiers like relative clauses, since such examples will invite the left dislocation parse with the null subject actually triggering plural agreement. Significantly, West and South Slavic languages allow null subjects (Franks [8]). See Sturgeon [11] for left dislocation in Czech.

Though genitive case marking of the head noun needs to be accounted for, I suggest that something else is directly responsible for neuter singular subject-verb agreement.

I will account for the behavior of higher numerals in the next section. Section 3 then compares higher numerals with paucals. Previous analyses are discussed in sections 3 and 4. Section 5 concludes.

2. A new proposal

What is it that forces singular agreement in (1) and its counterpart in Polish, Serbo-Croatian, and Slovene? My proposal is that the Slavic pattern should be assimilated to that of languages like Hungarian, where numerals in general require a singular noun and default singular subject-verb agreement, as in (3a).

(3) a. Hat gyerek elment.   Hungarian
         six   child.sg left.sg

     b. A gyerekek   elmentek.   (Farkas [12], p. 87)
        the child.pl   left.pl

Other languages where numerals force default agreement include Turkish and Georgian (Ortmann [13]). For these languages, one can say that [singular] is not specified with a value when a numeral is present.

I claim that essentially the same is true of Czech, Polish, Serbo-Croatian, and Slovene[2]. One complication, which I will get to in the next section, is that underspecification of [singular] is restricted to numerals 5 and above in these Slavic languages. Let us say that the numeral is merged into Spec of #P as in (4) and that as stated in (5), the # head which selects a higher numeral in Spec of #P cannot be specified with respect to [singular], which determines the shape of the head noun through agreement.

(4)   [#P numeral [# NP]]

(5)   Underspecification of [singular] in Czech, Polish, Serbo-Croatian, and Slovene
        The # head that selects higher numerals lacks the specification of [singular].

Thus, the parameter having to do with underspecification of [singular] resides in the lexicon.

There is evidence that the # head is not specified as [+singular]. As noted by Franks [4, 8], coordination of numerically quantified subjects requires neuter singular whereas coordination of ordinary singular neuter subjects results in masculine plural, as illustrated in (6) for Serbo-Croatian.[3],[4]

---

[2] It may not be a coincidence that these Slavic languages are spoken roughly in the same area as Hungarian and Turkish.

[3] The Serbo-Croatian quantifier *nekoliko* behaves in the same way as higher numerals with respect to subject-verb agreement (Wechsler and Zlatić [5], p. 120-121).

[4] The same pattern is found in Slovene (Marušič & and Nevins [6]), except that coordination of ordinary singulars triggers dual marking.

(6)   a. Pet devojaka i   nekoliko momaka je   skakalo.   Serbo-Croatian
          five   girl.gen.pl and   several boy.gen.pl aux.3sg   jumped.n.sg

          b. Tele i   dete   su   skakali.   Serbo-Croatian
          calf.n.sg and   child.n.sg   aux.3pl   jumped.m.pl (Franks [8], p. 115)

If feature resolution under coordination is based on specification in syntax, coordination of unspecified [singular] can only yield unspecified [singular], a correct prediction[5].

Though underspecification of [singular] in narrow syntax can explain why the presence of a higher numeral forces the default neuter singular agreement, it makes the genitive plural form of the head noun puzzling. Where does plural marking come from if [singular] is valueless? To solve this problem, we need to take into account the number feature system as a whole. I adopt the system in (7), which employs [±augmented] in addition to [±singular] to define the basic number distinction (see Harbour [15] and the references cited there).

(7)   The Number Feature System

      i.   singular: [+singular, –augmented]

      ii.   dual: [–singular, –augmented]

      iii.   plural: [–singular, +augmented]   (3)

Significantly, [+singular, +augmented] is an impossible combination under this system. This means that the minus value of [singular] in the feature makeup of plural is predictable given [+augmented]. And this is the conceptual basis for syntactic underspecification, which feeds semantic interpretation: predictable values can be omitted. Significantly, with numerals 5 and above, the number features must be specified as [–singular, +augmented], if full specification is used.

Let us now suppose that the predictable value of [singular], though absent in narrow syntax, is supplied in morphological computation. This value insertion can be implemented by (8).

(8)   [singular] —> [–singular] / [_____ , +augmented]

The background assumption here is that morphological computation is allowed to operate on the output of narrow syntax computation before morphosyntactic features receive phonological realization, as in Distributed Morphology (Halle and Marantz [16]).

One might wonder whether [±augmented] is motivated for languages like Czech, Polish, and Serbo-Croatian which lack the category of dual. Slovene preserves dual, but it was lost in Czech, Polish, and Serbo-Croatian. At this point, we can appeal to Watanabe's [17] proposal that [±augmented] is involved in the licensing of numerals. Specifically, he suggests (9) as a universal principle.

(9)   Numerals are licensed only when the # head is marked for [±augmented].

---

[5] For further intricacies of the agreement pattern triggered by coordinated subjects, see Bošković [14] on Serbo-Croatian and Marušič & and Nevins [6] on Slovene.

According to this hypothesis, [±augmented] is active in those languages which have numerals, even if dual is absent as an inflectional category. It follows that the use of [±augmented] is legitimate for languages like Czech, Polish, and Serbo-Croatian.

Incidentally, Watanabe [17] points out that the system in (7) allows [augmented] to be unspecified in the presence of [+singular], since the minus value is predictable in that case. Underspecification of [augmented], then, blocks the appearance of numeral 1, given (9). This phenomenon is fairly common with numerical bases, as documented by Hurford [18, 19]. Thus, the incompatibility of [+singular, +augmented] leads to two logically possible types of underspecification in the number feature system. For more details about the relation between [augmented] and numerals, see Watanabe [17].

The discussion so far has taken care of neuter singular subject-verb agreement and plural marking of the head noun. One remaining piece in the picture is genitive case. For this, I simply assume syncretism, of the form in (10i).

(10) i. structural case (nom, acc) —> gen / _____ [(+augmented),-less.than.a.handful]

    ii. oblique —> no change

(10i) is nothing more than a descriptive statement, and I leave vigorous formalization to future research. It will require an in-depth analysis of the entire case system in Slavic and more generally in Universal Grammar. Let me just mention that nominative, accusative, and genitive are implicated in other types of case syncretism in Slavic (see various chapters in Comrie and Corbett [20]), so that something like (10i) is a very plausible additional candidate.

(10ii) is intended to capture the fact that in Czech, Polish, and Slovene, both the numeral and the head noun exhibit the expected case form in oblique positions, as illustrated in (11) for Czech[6].

(11)   s   pěti   pány   Czech

       with five.inst   man.inst.pl   (Franks [8], p. 136)

It is worth reiterating at this point that it is not true that genitive marking of the head noun blocks agreement. In Polish predicate adjective constructions, the adjective itself appears in the genitive plural form, as in (12), despite the fact that the copular verb is neuter singular.

(12)   Szesc   kobiet   bylo   smutnych.   Polish

       six-nom woman.gen.pl be.past.n.sg sad.gen.pl   (Dziwirek [21], p. 147)

This means that the subject with a numeral is in principle capable of entering into the agreement relation. A peculiarity of Polish is that adjectives come with [±augmented] and [±less.than.a.handful] (in addition to [±singular]) as inflectional features, agreeing with the subject with respect to these two features. (8) and (10i) then ensure that the predicate adjective will take the genitive plural form. Verbs, on the other hand, lack

---

[6] In Serbo-Croatian, the head noun appears in the genitive plural form in oblique contexts, too, as discussed in Franks [4, 8] and Wechsler and Zlatić [5]. This difference seems to be correlated with the fact that the numeral is invariant in form irrespective of case in Serbo-Croatian, unlike in Czech, Polish, and Slovene. I will return to case forms of numerals below.

[±augmented] and [±less.than.a.handful] as well as case, preventing (8) and (10i) from applying[7].

The agreement in terms of [±augmented] and [±less.than.a.handful] is not limited to predicative adjectives in Polish. In the Czech example in (1), repeated here, the demonstrative and the attributive adjective take the form of genitive plural.

(1) Těch   pět hezkých   dívek   jelo.   Czech
      this.gen.pl five beautiful.gen.pl girl.gen.pl   traveled.n.sg   (Franks [8], p. 137)

This morphological shape also arises through (8) and (10i), which in turn are dependent on the working of agreement with respect to [±augmented] and [±less.than.a.handful].

3. Further consequences

As mentioned above, the default neuter singular subject-verb agreement is triggered by numerals 5 and above. The paucal numerals 2, 3, and 4 behave differently. The proposal outlined in the previous section enables us to understand why such a contrast exists.

Important for the purposes of this paper is the fact that the finite verb shows plural agreement when the subject contains a paucal numeral, as illustrated in (13) for Serbo-Croatian.

(13) a. Dva   srpska glumca   su   otišla / otišli.   Serbo-Croatian
         two   Serbian actor.m.234   aux.3pl   left.m.234/left.m.pl

     b. Dve   srpske glumice   su   otišle   Serbo-Croatian
         two   Serbian actress.f.234   aux.3pl left.f.pl(=f.234)

                                                      (Wechsler & Zlatić [5], p. 151)

Czech, Polish, and Slovene pattern in essentially the same way, except that Slovene uses dual for 2 (Corbett [7]). In other words, there is no syntactic underspecification for the paucal numerals. Higher numerals, on the other hand, force neuter singular, as shown in (14)[8].

(14) Pet ljudi   je   došlo   na   miting.   Serbo-Croatian
      Five people.gen.pl aux.3sg arrived.n.sg at   meeting   (Franks [8], p. 116)

This contrast between the paucal numeral and the higher numeral has not received a satisfactory account in the past. Franks [8] (p. 128, note 29) suggests that the adjectival status of the paucal numerals is responsible for plural agreement.[9] Though it is true that the Serbo-Croatian paucal numerals have three distinct case forms (nominative-accusative, genitive, oblique) unlike higher numerals, which are invariant, and that *dva*

---

[7] Given the existence of dual in Slovene verbs, further contextual conditions must be added to (8) and (10i), mentioning categorial information about applicable domains, to ensure the singular verb agreement. According to Corbett [22] (p. 134) and Marušič and Nevins [6] (note 1), the Slovene predicative adjective behaves in the same way as the Polish counterpart.

[8] Serbo-Croatian marginally allows plural, too (Franks [4, 8]; Wechsler and Zlatić [5]). I put this pattern aside, as essentially belonging to a different grammar.

[9] Veselovská [1] seems to follow suit.

'two' in addition is gender-sensitive, the account does not generalize to other languages[10]. Czech, for example, distinguishes two case forms for 5 and four case forms for 3, as shown in (15).

(15) Inflectional forms of Czech numerals (Veselovská [1])

　　　a. five: pět (nom, acc), pětk-I (gen, dat, loc, instr)

　　　b. three: tř-I (nom, acc, gen), tř-EM (dat), tř-ECH (loc), tře-MI (instr)

There is no sense in which three, but not five, is adjectival in nature. Besides, even if the adjectival status of the paucal numerals turns out to be relevant, it remains to be explained why the paucal numerals are adjectival, unlike higher numerals. The suggestion is nothing more than a restatement of the original problem[11].

Corbett [7], on the other hand, speculates that groups with a large cardinality "*are less individuated and are conversely more likely to be viewed as a unit*" (p. 217), leading to the nominal conception of the numeral in question and facilitaing singular agreement. To the extent that the categorial status is invoked, Corbett's proposal runs into the same problems as Franks'. Furthermore, the coordination data in (6) refutes the idea that singular agreement is triggered by the [+singular] feature. The failure of plural agreement in (6a) suggests that it is a mistake to attribute singular agreement forced by the subject with a higher numeral to the semantic notion of singularity, since coordination of semantically meaningful singulars yields plural agreement as in (6b). What is triggered by higher numerals is nothing more than default agreement.

There is further evidence that cardinality itself should not be blamed. Serbo-Croatian has two versions of *many*, one of them behaving like higher numerals and the other triggering plural agreement, as shown in (16).

(16) a. Mnogo　　　srpskih　　　pisaca　　　je otišlo.　　　Serbo-Croatian

　　　　many　　　　Serbian.gen.pl writer.gen.pl　aux.3sg　　　left.n.sg

　　　b. Mnogi　　　　　mladići　　　　　su　　　protestovali. Serbo-Croatian

　　　many.nom.m.pl young.man.nom.m.pl aux.3pl protested.m.pl

　　　　　　　　　　　　　　　　　　　　(Wechsler and Zlatić [5], pp. 116, 118)

As far as I am aware, there is no cardinality difference between the two versions of *many*. Plural agreement in (16b) shows that a large cardinality does not necessarily trigger singular agreement in this language. The idea of (non-)individuation is mistaken, too. Franks [4] (p. 626) [8] (p. 116) observes that both group and individuated readings are available for (14). Thus, it is an error to associate the idea of less individuated group members with a large cardinality to account for the pattern of subject-verb agreement. What is going on, then?

---

[10] Wechsler and Zlatić [5] (p. 149) remark that the genitive and oblique forms are in fact not used with 3 and 4, which are essentially frozen. This denies the adjectival nature of these two numerals.

[11] It should also be noted that Polish higher numerals change their shape in agreement with the gender information of the head noun in structural case contexts. Their inflection is therefore "adjectival". See Franks [4, 8] and Rappaport [2] for discussion.

I claim that underspecification holds the key, again. An important auxiliary hypothesis is that syntactic underspecification is enforced in a set of contexts provided independently. In Slavic, [±less.than.a.handful] divides plural numerals into two classes. Significantly, this division affects the case form of the head noun, in addition to the value of [±singular], the target of underspecification. In Serbo-Croatian, the paucal numerals force the head noun to take what Browne [23] calls the 234 form, which is syncretic with genitive singular for masculine and neuter nouns and with nominative plural for feminine nouns (Franks [8], p. 125, note 8). Higher numerals, on the other hand, require genitive plural, as discussed above. The relevant forms of a masculine noun are given in (17).

(17) a. prózori (nom.pl)   'window'   Serbo-Croatian

      b. prózora (gen.pl)

      c. prózora (234 = gen.sg)   (Browne [22], p. 319)

Thus, the division would be needed even if these languages did not choose underspecification of [±singular][12].

A similar classification of numerals is found in Irish, quite independently of under-specification. The phenomenon sensitive to the division is initial mutation affecting the head noun that follows the numeral. The dividing line here is between 6 and 7. Lenition applies in (18a), while we find nasalization in (18b).

(18) a. dhá/trí/ceithre/cúig/sé   chat   (lenition)   Irish

          two/three/four/five/six   cat.sg

      b. seacht/ocht/naoi/deich   gcat   (nasalization)

          seven/eight/nine/ten   cat.sg   (Acquaviva [24], pp. 165, 167)

The head noun appears in the singular form in both cases, and thus does not vary in number marking. See Acquaviva [24] and the references cited there for further discussion of Irish numerals.

We can conclude that the numeral system can introduce within itself a division that affects the morphosyntax of the head noun in various ways as a parametric option. In the relevant Slavic languages, I claim that case is the primary factor in this division, as encoded by whatever is the correct analysis of syncretism described by (10i). Underspecification of [±singular] chooses a subclass of numerals already given in terms of case considerations. It then follows that the subclass in question must be [–less.than.a.handful], since syntactic underspecification is based on predictability of the unspecified value. Crucially, in the class defined by [+less.than.a.handful], the value of [singular] is not always predictable from the value of [augmented]. 2 is specified as [–singular, –augmented], but [–singular] is not predicted by [–augmented], which is

---

[12] In Czech and Polish, the head noun after the paucal numerals appears in the form required by an external case assigner. Hence nominative plural in the subject position, though with some twists for masculine human nouns in Polish (Rappaport [2]). Genitive plural after higher numerals, on the other hand, is a common Slavic trait. The division thus affects case forms in languages other than Serbo-Croatian as well.

compatible with [+singular] as well, as far as combinatorial possibilities of feature values are concerned. Therefore, the class of [+less.than.a.handful] cannot be chosen as the domain of underspecification.

Note that this result hinges on the assumption that the predictability of the [±singular] value comes from [+augmented], but not from the numeral itself. If the numeral itself acts as the predictor of the unspecified value, underspecification can be enforced anywhere. And this latter possibility is found in languages like Hungarian, where any numeral forces the head noun to be in the singular. We then have the following typology of syntactic underspecification of [±singular] in the context of the numeral:[13]

(19) Underspecification of [±singular] in the context of the numeral

	predictor	domain
i. Hungarian Type	numeral	all
ii. Slavic Type	[+augmented]	[–less.than.a.handful]

These two must be the basic types. When there is a numeral, the value of [singular] can be predicted either by [+augmented] or by the meaning of the numeral itself. No other possibility exists. If [+augmented] is the predictor, the entire numeral domain cannot be selected, since singular and dual are [–augmented]. There must be a class of higher numerals independently given by other morpho-syntactic considerations. If the meaning of the numeral itself matters, the entire domain can be chosen. An open empirical question is whether a subclass of numerals can also be picked out as the domain of underspecification, arbitrarily this time, when the numeral acts as the predictor of the [±singular] value. So far, no such case is reported. If this third type does not exist, we can say that selection of the domain of underspecification is maximal.

There is another point to be made. As proposed above, the Slavic underspecification is accompanied by value insertion during morphological computation. I suggest that this value insertion is possible because the predictor of the value is [+augmented], a feature available to morphological operations. In other words, the missing value can be supplied in morphology only when that value is predicted by another feature available in morphology.[14] On the other hand, the semantic content of the numeral cannot play a role in morphology. It is simply sent to the LF interface for semantic interpretation. Thus, it is predicted that the Hungarian type underspecification will disallow value insertion in morphology, hence always visible in a transparent way.

4. Comparison with previous analyses

---

[13] I put aside vague quantifiers like *many* here, though they also trigger underspecification in Slavic, as we have seen in (16). They belong to a separate system of quantification. See Watanabe [17] for arguments that they are structurally different from numerals.

[14] In order to make this idea work, one probably has to say that the predictor feature (value) is marked as such, hooked to (8), so that its predictor status is visible to the morphological component. I leave it to future research to explore implications of this mechanism.

Let us take stock. A subset of Slavic languages underspecify [singular], based on the fact that [+augmented] must be combined with [–singular]. The underspecification is associated with the # head that selects a class of numerals 5 and above, a division which must be independently provided for case morphology. This association is forced by the choice of [+augmented] as the predictor of the missing value of [singular]. The underspecification in narrow syntax, however, is masked by "repair" in morphology.

Above, we have already seen that the previous analyses have not succeeded in explaining why singular agreement is forced by higher numerals, but not by the paucal numerals. Let us now focus on the part of the accounts that deals with the forced default singular agreement itself in the case of higher numerals.

Franks [8] claims that the plural feature is blocked from percolating to the top projection by oblique genitive marking and hence inaccessible for agreement in Serbo-Croatian, whereas higher numerals are always in the accusative in Czech and Polish, failing to induce agreement. It is highly problematic that a rather intricate common set of agreement problems does not receive a unifying account. Furthermore, we have seen that subject-predicate agreement is not completely blocked, as evidenced by Polish data. Recall that Polish predicate adjectives are inflected as genitive plural in (12). This fact can be accounted for by (8) and (10i) if adjectives have [+augmented] and [–less.than.a.handful]. The value of these features must come from agreement with the subject.

Wechsler and Zlatic [5] propose for Serbo-Croatian that higher numerals lack phi-features and therefore lead to default agreement[15]. Polish data on predicate adjectives, again, indicate that default agreement is used for [singular], but not for [augmented].

5. Conclusion

In this paper, I have explored the idea that underspecification in narrow syntax is responsible for singular subject-verb agreement induced by numerals 5 and above in Czech, Polish, Serbo-Croatian, and Slovene. This is an instance of disguised underspecification, since morphology repairs the underspecification of [singular] in narrow syntax. A novel result is that we now have a principled explanation of why the singular agreement is forced by higher numerals but not by the paucal numerals. This explanation is made possible by the very nature of underspepcification, which must be motivated by the predictability of the missing feature value. The predictor of the missing feature value is parametrized. In the relevant Slavic languages, it is keyed to [+augmented].

What masks syntactic underspecification is a morphological operation that provides the predictable feature value missing in narrow syntax. This morphological operation must be distinguished from the insertion of the unmarked value advocated by Harbour [25] and Noyer [26]. Quite generally, disguised underspecification can be regarded as a

---

[15] Veselovská [1], p. 302 seems to assume something similar for Polish.

major source of agreement mismatch. It is an interesting task for future research to apply the idea to various individual cases.

References

[1] L. Veselovská, Agreement Patterns of Czech Group Nouns and Quantifiers, in Semi-Lexical Categories, eds. N. Corver and H. van Riemsdijk, Mouton de Gruyter, Berlin, 2001, pp. 273–320.

[2] G. Rappapor, Case Syncretism, Features, and the Morphosyntax of Polish Numeral Phrases, in Generative linguistics in Poland 5, 2003 pp. 123–137.

[3] P. Rutkowski and H. Maliszewska. On Prepositional Phrases inside Numeral Expressions in Polish, Lingua 117, 2007, pp. 784–813.

[4] S. Franks, Parametric Properties of Numeral Phrases in Slavic, Natural Language and Linguistic Theory 12, 1994, pp. 597–674

[5] Wechsler, S. and L. Zlatić, The Many Faces of Agreement, CSLI Publications, Stanford, 2003.

[6] F. Marušič and A. Nevins, Two Types of Neuter: Closest-Conjunct Agreement in the Presence of '5 and Ups', paper presented at Formal Approaches to Slavic Linguistics 18, 2009.

[7] Corbett, G. G., Number, Cambridge University Press, Cambridge, 2000.

[8] Franks, S., Parameters of Slavic Morphosyntax, Oxford University Press, Oxford, 1995.

[9] Ž. Bošković, Case and Agreement with Genitive of Quantification in Russian, in Agreement System, ed. C. Boeckx, John Benjamins, Amsterdam, 2006, pp. 99–120.

[10] A. Pereltsvaig, Small Nominals. Natural Language and Linguistic Theory 24, 2006, pp. 433–500.

[11] Sturgeon, A., The Left Periphery: The Interaction of Syntax, Pragmatics and Prosody in Czech, John Bejmains, Amsterdam, 2008.

[12] D. F. Farkas, The Unmarked Determiner, in Non-Definiteness and Plurality, eds. S. Vogeleer and L. Tasmowski, John Benjamins, Amsterdam, 2006, pp. 81–105.

[13] A. Ortmann, Where Plural Refuses to Agree: Feature Unification and Morphological Economy, Acta Linguistica Hungarica 47, 2000, pp. 249–288.

[14] Ž. Bošković, Unifying First and Last Conjunct Agreement, Natural Language and Linguistic Theory 27, 2009, pp. 455–496.

[15] Harbour, D., Morphosemantic Number, Springer, Dordrecht, 2007.

[16] M. Halle and A. Marantz, Distributed Morphology and Pieces of Inflection, in The View from Building 20, eds. K. Hale and S. J. Keyser, MIT Press, Cambridge, 1993, p. 111–176.

[17] A. Watanabe, Vague Quantity, Numerals, and Natural Numbers, *Syntax* 13, 2010, p 37-77.

[18] Hurford, J. R., Language and Number, Basil Blackwell, Oxford, 1987.

[19] J. R. Hurford, The Interaction between Numerals and Nouns, in Noun Phrase Structure in the Languages of Europe, ed. F. Plank, Mouton de Gruyter, Berlin, 2003, pp. 561–620.

[20] Comrie, B., and G. G. Corbett, eds., The Slavonic Languages, Routledge, London, 1993.

[21] K. Dziwirek, Default Agreement in Polish, in Grammatical Relations, eds. K. Dziwirek, P. Farrel, and E. Majías-Bikandi, CSLI Publications, Stanford, 1990, pp. 147–161.

[22] Corbett, G. G., Agreement, Cambridge University Press, Cambridge, 2006.

[23] W. Browne, Serbo-Croat, in The Slavonic Languages, eds. B. Comrie and G. G. Corbett, Routledge, London, 1993, pp. 306–387.

[24] Acquaviva, P., Lexical Plurals, Oxford University Press, Oxford, 2008.

[25] D. Harbour, The Kiowa Case for Feature Insertion, Natural Language and Linguistic Theory 21, 2003, pp. 543–578

[26] R. Noyer, Impoverishment Theory and Morphosyntactic Markedness, in Morphology and its relation to phonology and syntax, eds. S. G. Lapointe, D. K. Brentari, and P. M. Farrell, CSLI Publications, Stanford, 1998, pp. 264–285.

Morphological productivity and family size:
evidence from French compound nouns garde-x and N-de-N

ELENA VOSKOVSKAIA
University of Toronto
helen.voskovskaia@utoronto.ca

0. Abstract

Recent work in derivational morphology have shown a new interest for the concept of morphological productivity and its various measuring tools, both quantitative (Baayen, 1992; 1993; Baayen & Lieber, 1991; Baayen & Renouf, 1996) and qualitative (Bauer, 2001; Dal, 2003). However, the productivity of compound words has been very little explored, especially in French, except for the work of Krott, Schreider & Baayen (1999). Recent studies have proposed a new approach to morphological productivity in terms of family size (De Jong, Schreuder & Baayen, 2000; del Prado M., Bertram, et al., 2004). This paper explores a possible correlation between morphological productivity and family size and is based on the results of a quantitative study of French compounds garde-x and N-de-N from 1606-1920. The result of study indicates that there is a negative correlation between the productivity and family size of compounds garde-x. The growth of the morphological family increases the level of productivity of these compounds. In case on N-de-N compounds, the study demonstrated mixed results: the correlation was negative for first and second analysed periods, but positive for the third and fourth period.

0. Résumé

Des travaux récents dans la morphologie dérivationnelle ont démontré un nouvel intérêt pour le concept de la productivité morphologique et pour ses outils de mesure en utilisant une méthode quantitative (Baayen, 1992; 1993; Baayen & Lieber, 1991; Baayen & Renouf, 1996) et qualitative (Bauer, 2001; Dal, 2003). Cependant, la productivité des mots composés a été très peu explorée, particulièrement en français, à part le travail de Krott, Schreider et Baayen (1999). Récemment, une nouvelle approche vers la mesure de la productivité morphologique en terme de la taille de la famille morphologique a été proposée (de Jong, Schreuder & Baayen, 2000; del Prado et al., 2004). Ce papier étudie une corrélation possible entre la productivité morphologique et la taille de la famille morphologique des composés français garde-x et N-de-N des 1606 à 1920. Le résultat de la recherche indique qu'il existe une corrélation négative entre la productivité et la taille catégorielle de la famille morphologique des composés garde-x. La croissance de la famille morphologique augmente le niveau de la productivité de ce type de composés. En ce qui concerne les composés N-de-N, la recherche a montré des résultats mixtes : la corrélation était négative pour la

première et deuxième période étudiée, mais elle était positive pendant la troisième et quatrième période.

Keywords: Morphology, morphological productivity, morphological family size, diachrony, linguistic corpus, French compounds

Mots-clefs : Morphologie, la productivité morphologique, la taille de la famille morphologique, la diachronie, le corpus linguistique, les composés français

1. Introduction

The notion of productivity can be approached from both a qualitative or quantitative angles. From qualitative standpoint, the productivity is defined as the capacity of morphological rules to form new lexemes in a non intentional manner (Dal 2003: 5) using the available morphological process (Bauer 2001:49). From the quantitative standpoint, several methods based on the linguistic corpora were proposed to measure different aspects of productivity (Baayen & Lieber, 1991; Baayen, 1992; 1993; Baayen & Renouf, 1996). The productivity of French compound words is very little explored: except for the work of Krott, Schreider & Baayen (1999) and Fernández-Domínguez (2007, 2008) the studies on the measure of productivity have not been applied to compounds.

Currently, the often used calculations of morphological productivity are presented by Baayen (1991, 1996, 1999). Baayen proposes two measures of morphological productivity: 'the productivity in the strict sense' and the 'global productivity' - the two measures are calculated comparatively to a big corpus. The global productivity, noted P*, expresses the probability of appearance of a new word in a given morphological process (Hay & Baayen, 2002).[1] The method more used in the present research on the morphological productivity is a statistical measure in the strict sense (Baayen & Lieber, 1991; Baayen, 1992).The productivity in the strict sense, noted as *P*, represents a quantitative measure that is based on the notion of hapax legomenon.

Besides a traditional quantitative analysis of the morphological productivity, another approach to the measure of productivity was recently rediscovered in the linguistic environments – the correlation between the morphological productivity and the size of the morphological family. Prado *et al.* (2005:496) show that the size of the morphological family is strongly correlated to the word frequency: if a word is frequent, the size of its morphological family tends to be bigger. The *categorial family size* is a measure showing how the entire morphological category contributes to morphological connectivity in the mental lexical process.

While analyzing the correlation between the productivity of the affixes and the size of its morphological family in English, Baayen and Hay (2002) show that the increase of the

---

[1] Baayen represents in his first works (1989, 1991) the global productivity as a possibility to incorporate the profitability of the studied process in the measure of the productivity (two process A and B can have the same P value while the one forms ten times more words than the other).

categorial family size causes growth of its morphological productivity. They define the categorial family size of the morphological family as following: « …*the 'categorial family size' of a morphological category be the sum of the (logarithms of the) family sizes of the base words of the members of this category*" (Baayen & Hay, 2002:1). For example, the categorial family size of the affix *un-* in the words *unbound, unsafe, unbrave, unwise* is a sum *log (F1) + log (F 2) + log (F 3) + log (F4)* etc. where F1, F2, F3 represents the family size of the word-bases *bound, safe, brave, wise*, etc.

The measure of the morphological productivity, very little explored for compound words, raises an important question*: Is there evidence of a correlation between the productivity and size of the morphological family of the compounds?*

The aim of this paper is to verify the potential correlation between the productivity of French compounds and the categorial family size of their morphological families. Our hypothesis is the following one: There is a positive correlation between the productivity and size of the morphological family of the compounds *garde-x* and N-de-N: the growth of the categorial size of its morphological family increases the productivity level of compounds.

Using Bayeen's (1991, 1992) productivity measure *P* in the strict sense and the model of the categorial family size (Baayen & Hay, 2002), this paper examines the productivity of the French compound nouns and explores the potential correlation between the morphological productivity and categorial size of its morphological family.

The research is based on French corpus *Frantext* through the analysis of 60 compounds *garde-x* and 49 compounds N-de-N.

Following the introduction, section 2 presents the methods of morphological productivity measurement used in present research. The corpus and analysed data are described in section 3. Section 4 discusses the results of research on the productivity of French compounds *garde-x* and N-de-N and the possible correlation between the productivity and categorial family size. Sections 5 and 6 outline major discussion and conclusions.

2. Methods of productivity measurement

In our study, two measures were used to evaluate the productivity of French compounds *garde-x* and N-de-N: the measure *P* based on the notion of hapax legomenon (Baayen & Lieber 1991: 809; Baayen 1992: 115-16; Hay & Baayen 2003: 101) and the measure of the categorial family size (Baayen & Hay : 2002). Evaluation of productivity based on the hapaxes considers the productivity of the morphological process inside the same corpus (Frantex) and it is calculated using the following formula:

(1) $P = n_1 / N$

where P= productivity; n_1 = the number of hapax legomenon (the forms that appear only once in a big corpus); N = the total token frequency of words created in the corpus. This calculation evaluates the type productivity of compounds as a whole structure while showing the rate of growth of the vocabulary.

According to this measure of productivity, a category with a big number of words of high frequency will have a big value of N and, consequently, a lower degree of productivity. This calculation evaluates the type productivity of compound nouns as a fixed (whole) structure while showing the rate of growth of the vocabulary. In order for the hapaxes to represent true neologisms, it is necessary that the research of the *P* is done on a very big corpus (tens of millions of words).

The second measure implied in the evaluation of the morphological productivity of compounds *garde-x* and N-de-N is the measure based on the categorial family size of the morphological family proposed by Baayen and Hay (2002). In their work on the productivity of the English affixes and the bases of derived words, Baayen and Hay (2002:1) show that there is a strong correlation between the degree of productivity of the base and the size of the morphological family. Words such as *man* or *coin* are presented as constituents (bases) in several complex words; they possess a high degree of productivity and their morphological families are quite big. This effect, linked to the size of the morphological family, was also observed in Dutch, German, English, Hebrew and Finnish.

The measure of the categorial family size of the morphological family proposed by Baayen and Hay (2002) is calculated in two steps.

a) First, the size of the morphological family of the analysed form is measured (for example, the family of the word *chou-fleur*, includes all the derived and compound words formed from this base, except the bases themselves: *chouchou, chouchouter, chou-blanc, fleurir, fleurette, fleuriste, fleuron*, etc). According to the definition of Schreuder and Baayen (1997: 118), the morphological family includes all the words that are formed with the bases word through derivation or composition.

b) Secondly, to measure a categorial family size of a morphological family, a logarithmic formula representing the sum of the (logarithms of the) family sizes of the base words of the members of this category is applied:

(2)   log (F1) + log (F 2) + log (F 3) etc.

where F1, F2, F3, etc. are the sizes of the morphological family of the base words[2]. The measure of the categorial family size calculates the frequency of the constituent of the compounds as well as the frequency of the members of their morphological families.

3. Corpus and material preparation

Since the measure of morphological productivity based on hapax legomenon requires that forms are extracted from a large corporus, two sources were used to create a primary database of French compound nouns. The Gallica corpus of French National Library (50 million words) was employed to develop a list of 60 compounds *garde-x*.

---

[2] Following Schreuder et Baayen (1997), Moscoso del Prado et al. also suggest that the effect of the size of the family is logarithmic by the nature: "Like the word-frequency effect, the family-size effect is logarithmic in nature. Robust effects are typically observed in the range of 0–40 family members, after which there is generally a floor effect". (Moscoso del Prado et al. 2004 :1272)

Littré's *Dictionnaire de la langue française* (1863-1877) with 80,000 entries was used to create a list of 49 compounds N-de-N. Gallica corpus of the National Library of France offers 1200 volumes in mode text that represent various domains such as history, literature, science, philosophy, law, economics and political science. This corpus reflects a general portrait of the French language usage from 16th to 21st centuries. Littré's *Dictionnaire* is a normative dictionary that reflects a state of the classical French language and its good literary usage between the 17th and the 19th centuries; this dictionary remains an important stage in the development of the French language.

The larger corpus Frantext[3] was later used to validate the frequency of each token and type of the words being analysed. Frantext database consists of up to 4000 texts (more than 150 million words) covering several domains such as science, arts, literature, technology, politics, philosophy, etc. This corpus has a sufficient size to apply a series of quantitative measures and analyse necessary statistics such as token number, hapaxes, frequency of family members, etc.

The corpus is divided in four periods according to their importance in the history of the French language: 1606-1694 (17.3 million words); 1695-1798 (34.4 million words); 1799-1872 (41 million words); 1873-1920 (28 million words). The hapaxes that appear in Frantext are good markers of the productivity of the studied process (in other terms, their aptitude to form new words): the bigger is the number of hapaxes, the more productive is a studied morphological process. This is the reason why the size of corpus is extremely important: a too limited corpus risks misrepresenting the unique forms such as neologisms.

The *garde-x* compounds analysed in the current research are distributed over the following types:

(3)   a. N-N : garde-temps, garde-sel, garde-notes

        b. V-N : garde-main, garde-cote, garde-feu

        c. N-Prép-N: garde de santé, garde de chasse

        d. N-A : garde-fou

N-de-N compounds represent the following types:

(4)   a. bec-de-N : bec-de-lézard , bec-de-cigogne, bec-de-cane, etc.

        b. belle-N : belle-de-jour, belle-de-nuit, etc.

        c. cul-de-N: cul-de-jatte, cul-de-poule, cul-de-sac, etc.

        d. coup-de-N : cou-de-pied, coup de poing, etc.

---

[3] The Frantext database (http://humanities.uchicago.edu/orgs/ARTFL/ (formerly the *Trésor de la Langue Française*) was elaborated in 1960. It included less than thousand works from 19th and 20th centuries selected from the bibliographies of the history of literature. This corpus quickly became enriched towards 1970 with scientific and technical texts. Afterward, the funds spread in a manner to cover the three preceding centuries as well as the following period. The corpus continues to grow progressively.

Table 1 gives a general overview of the statistics representing the total token frequency of compounds *garde-x* and N-de-N in Frantext from 1606 to 1920:

TABLE 1: Token frequency of compounds N-de-N and garde-x in Frantext corpus

	1606-1694	1695-1798	1799-1872	1873-1920	Total
N-de-N	113	173	204	277	767
garde-x	118	320	502	147	1087
works	141	183	283	64	671
words	4,884,725	12,524,257	20,608,307	5,208,641	43,225,930

The results included in Table 1 show that N-de-N compounds were used more frequently during the third and fourth periods (204 and 277 occurrences respectively); *garde-x* compounds demonstrated a higher frequency during the second and third studied periods (320 and 502 respectively).

4. Results

4.1 Morphological productivity of garde-x compounds During research, 25 hapaxes related to compounds *garde-x* were observed. While applying the measure based on hapaxes of Baayen (Baayen & Lieber, 1991; Baayen, 1992), the morphological productivity of the compound nouns *garde-x* was calculated as follows:

$$(5) \quad P = \frac{25}{1316} = 0.019$$

where 25 is the number of forms *garde-x* formed by the compounding process and occurring in Frantext only once; 1316 is the total token frequency of compounds *garde-x*.

The changes in the rate of the morphological productivity of the compounds *garde-x* showed the following results summarized in Table 2:

TABLE 2: Productivity of compounds *garde-x* in Frantext corpus

Période	n_1	N	P
1606-1694	2	75	0.027
1695-1798	8	430	0.019
1799-1872	2	491	0.004
1873-1920	13	320	0.041
Total :	25	1316	0.127

where n_1 is the number of forms *garde-x* formed by the compounding process and occurring in Frantext only once; N is the total token frequency of compounds *garde-x*; P is the morphological productivity of compounds *garde-x*.

Table 2 shows that the morphological productivity of the compounds *garde-x* was elevated during the fourth and the first periods ((P=0.041 and P=0.027 respectively). The rather high level of productivity of this type of compounds during the first period could be explained by loss of importance of Latin and the introduction of French language into daily life through several neologisms. The morphological productivity of compounds *garde-x* significantly lowered during the years 1799-1872 (P=0.004).

4.2 Productivity and categorial family size of compounds garde-x

According to the definition of Schreuder et Baayen (1997: 118), the morphological family includes all the words that are formed with the word base through derivation or compounding, while excluding the base itself.

To measure the categorical family size of the compounds *garde-x*, we followed the two steps mentioned above:

1) First, the size of the morphological family of every compound was calculated for each of the four periods. To measure the size of the morphological family of compounds *garde-x*, we released all the words belonging to the same morphological family of each of *garde-x* form constituent. For example, for the compound *garde-chasse*, we were looking in Frantext for the words such as *garder, gardeur, gardeuse, gardien, gardienne,* etc. (garde) and *chasser, chasseur, chasseresse,* etc. (chasse).

The distribution of the morphological family members of the compounds *garde-x* by period is presented in Table 3:

TABLE 3: Morphological family size of compounds *garde-x* in Frantext corpus

garde-x	1606-1694	1695-1798	1799-1872	1873-1920
types of compound	9	20	27	16
size of morphological family (total)	46	134	231	200
size of morphological family (average)	5.1	6.7	8.6	12.5

While analyzing the data in Table 3, we noticed a significant growth in the average number of the size of morphological members for four studied periods: from 5,1 members fin the first period to 12,5 members in the fourth period.

2) Secondly, the logarithmic formula was applied to calculate the categorical family size of the compounds *garde-x*. For example, the categorical family size of this type of compound for the first period was measured as follows:

(6) $\log(6) + \log(8) + \log(5) + \log(4) +$ etc. $= 6.23$

where $\log(6)$, $\log(8)$, $\log(5)$, $\log(4)$, etc. are the logarithms of the family sizes of the base words.

The correlation between the level of productivity and the categorical family size of compounds *garde-x* is captured in the Table 4:

TABLE 4:Morphological productivity and family size of compounds *garde-x* in Frantext corpus

Mesure	1606-1694	1695-1798	1799-1872	1873-1920	average
N	75	430	491	320	329
F	46	134	231	200	152.75
T	6.23	16.43	24.87	15.80	15.8
P	0.027	0.019	0.004	0.041	0.023

where N is the total token frequency of compounds *garde-x*; F is the size of the morphological family; T is categorical family size of the morphological family; and P is the morphological productivity of compounds *garde-x*.

The analysis of the data in the Table 4 shows that in the case of compounds *garde-x*, there is a negative correlation between the morphological productivity and the specific category size of the morphological family. The compounds related to the third period (1799-1872) possessing the highest categorical family size (T=24.87) have the lowest level of productivity (P=0.004). The compounds with the smallest categorical family size ((T=6.23) possess a higher level of productivity (P=0.027).

4.3 Morphological productivity of N-de-N compounds

According to the measure based on the number of words of the given category that occur only once in the corpus (Baayen & Lieber, 1991; Baayen,1992), 23 hapaxes were discovered for compounds N-de-N in Frantex corpus (all periods combined). The morphological productivity of the compound N-de-N was calculated as follows:

$$(7) \quad P = \frac{23}{767} = 0.030$$

The productivity of compounds N-de-N differs across different studied periods. The results of our research are summarized in Table 5.

TABLE 5: Productivity of compounds *N-de-N* in Frantext corpus

Période	n_1	N	P
1606-1694	5	113	0.044
1695-1798	3	173	0.017
1799-1872	6	204	0.029
1873-1920	9	277	0.032
Total :	23	767	0.122

The analysis of Table 5 permits to conclude the following:

1. The compound N-de-N were more productive during the period from 1606 to 1694 (P=0.044).

2. The level of the morphological productivity of compounds N-de-N was 2.5 times lower (P=0.017) during the years 1695-1798 in comparison with the first period.

4.4. Productivity and categorial family size of compounds N-de-N

The categorical family size of the compounds N-de-N was calculated following the same two steps mentioned above for compounds *garde-x*.

   1) First, the size of the morphological family of compound N-de-N was determined for each of the four periods as F1=63; F2=70; F3=90 and F4=102 respectively.

   2) Secondly, the logarithmic formula was applied. For example, for the first period (1606-1694), the categorical family size of compounds N-de-N with 12 different types was measured as follows:

   (8)   log(4)+log(3)+log(6)+log(6)…..=8.17

The correlation between the rate of productivity and the categorical family size of the compounds N-de-N is outlined in Table 6.

TABLE 6: Morphological productivity and family size of compounds N-de-N in Frantext corpus

mesure	1606-1694	1695-1798	1799-1872	1873-1920	average
N	113	173	204	277	192
F	63	70	90	102	81
T	8.17	8.42	11.66	13.56	10.45
P	0.044	0.017	0.029	0.032	0.031

where N is the total token frequency of compounds N-de-N; F is the size of the morphological family; T is categorical size of the morphological family; P is the morphological productivity of compounds N-de-N.

Table 6 represents the vocabulary growth for compounds N-de-N throughout four studied periods. Our results can be summarized as follows. First, we have seen a negative correlation between the productivity and the categorical family size of compounds N-de-N through two studied periods (P1=0.044; P3= 0.029 vs T1=8.17; T2=8.42 respectively). Secondly and conversely, instead of lower productivity level for bigger categorical family size as for *garde-x* compounds, we observe the positive correlation between the morphological productivity and family size of compounds N-de-N for the third (1799-1872) and fourth period (1873-1920): P3=0.29; P4=0.032 vs T3=11.66; T4=13.56 respectively.

5. Discussion

The study conducted on the compound nouns *garde-x* and N-de-N evaluated the quantitative productivity of these forms using two different statistical measures: the

measure based on the hapaxes ((Baayen & Lieber, 1991; Baayen,1992; Hay & Baayen, 2003) and the measure of the family size of the morphological category ((Baayen & Hay, 2002).

This study showed that compounds *garde-x* and N-de-N can be considered productive in the analyzed corpus: the average productivity of compounds *garde-x* through four analyzed periods is P=0.023; the compounds N-de-N demonstrate an average productivity during the same periods as P=0.031.

The aim of this study was to verify the possible correlation between two measures of productivity for compounds *garde-x* and N-de-N. We advanced the hypothesis that there is a positive correlation between the productivity and size of the morphological family of the compounds *garde-x* and N-de-N: the growth of the categorial size of its morphological family would increase their level of productivity (following the discovery of positive correlation between the productivity of the affixes and the size of its morphological family in English by Baayen and Hay (2002)).

The results presented in the Table 4 and 6 show that, globally, in the matter of compounds *garde-x* and N-de-N, there is a negative correlation between the productivity and the categorial size of their morphological families. For *garde-x* compounds, this correlation is evident for the first (1606-1694), second (1695-1798) and third period (1799-1872) where the growth of the family size of a morphological category (T1=6.23; T2=16.43; T3=24.87) diminishes the level of productivity of this type of compounds (P1=0.027; P2=0.019; P3=0.004). For N-de-N compounds, this correlation stays negative for the first and second period (T1=8.17; T2=8.42 vs P1= 0.044 and P2= 0.017 respectively).

It is interesting to note that the negative correlation between morphological productivity and categorial family size does not hold for compounds *garde-x* during the fourth period (T4=15.80 vs P4=0.41). The data shows that the highest level of productivity is related to a medium categorial family size of the morphological family for this type of compounds. In the matter of compounds N-de-N, this correlation was positive for the third and fourth analysed periods (T3=11.66 vs P3=0.029 and T4=13.56 vs P4=0.032). At the same time, the highest level of productivity of compounds N-de-N is correlated to the smallest categorial size of their morphological family (T1=8.17 vs p1=0.044).

6. Conclusion

In contrast to the discovery of a positive correlation between the productivity of the affixes and the size of their morphological family in English (presented by Baayen and Hay (2002)), the results of the study of the compound nouns *garde-x* and N-de-N show that this correlation does not always hold for French compounds. Generally, the growth in the categorial size of the morphological family of the compound *garde-x* and N-de-N decrease the level of their morphological productivity.

The negative correlation between the level of productivity and the size of the morphological family for the compounds *garde-x* and N-de-N during the first (1606-1694) and the second periods (1695-1798) can be explained by the specific linguistic situation in France at that period of time. The fact that the Latin was less frequent and began losing its importance at this time allowed the French language to form new words and be widely introduced into daily life through several neologisms.

Our findings suggest that more research on different type of French compounds and their morphological families is necessary.

References

Baayen, R. H.,Lieber, R. 1991. Productivity and English derivation: a corpus-based study, Linguistics, 29, pp. 801-43.

Baayen, R.H. 1992. Quantitative aspects of morphological productivity, Yearbook of Morphology, Kluwer Academic Publishers, , pp.109-149.

Baayen, R.H. 1993. On frequency, transparency and productivity, Yearbook of Morphology, Kluwer Academic Publishers, pp.181-208.

Baayen, H., Renouf A. 1996. Chronicling the Times: productive lexical innovations in an English newspaper, Language 72: 69-96.

Bauer, L. 2001. Morphological productivity, Cambridge: Cambridge University Press.

Dal, G. 2003. Productivité morphologique : définitions et notions connexes, dans Langue française, v.140, Cedex, Larousse, Paris, pp.3-23.

De Jong, N. H., Schreuder, R. and Baayen, R. H. 2000. The morphological family size effect and morphology, Language and Cognitive Processes 15, 329–365.

De Jong, N.H., Schreuder, R., and Baayen, R.H. 2003. Morphological resonance in the

mental lexicon. In Trends in Linguistics: Morphological Structure in Language Processing, eds.,

R.H. Baayen and R. Schreuder, 65-88. Berlin, Germany: Mouton de Gruyter.

del Prado, M., Bertram, R., Haikio, T., Schreuder, R., Baayen, H. 2004. Morphological Family Size in a Morphologically Rich Language: The Case of Finnish Compared With Dutch and Hebrew, Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 30, No. 6, p.1271–1278.

Dijkstra, T.; del Prado, M.; Schulpen, B.; Schreuder, R.; Baayen, H.; 2005, A roommate in cream: Morphological family size effects on interlingual homograph recognition, in: *Language and Cognitive Processes*, Volume 20, Issue 1 & 2 , pages 7 – 41.

Fernández-Domínguez, J., Diaz-Negrillo, A., Pavol Štekauer, P., 2007. **How is Low Morphological Productivity Measured?.** *Atlantis* 29.1: 29-54.

Fernández-Domínguez, J., 2009. Productivity in English Word-formation. An approach to N+N compounding, Eurppean University Studies, Peter Lang.

Hay, J., Baayen, H., 2002. Affix Productivity and Base Productivity. Paper presented at the Morphological Productivity Seminar, ESSE 6, Aug 30-Sept 3, Strasbourg.

Krott, A., Schreuder, R., Baayen, R. 1999. Complex words in complex words, Linguistics, 37-5, pp. 905-926.

Schreuder, R. and Baayen, R. H.1997. How complex simplex words can be, Journal of Memory and Language 37, 118–139.

Les corpus

Frantext : University of Chicago. ARTFL Project.

http://humanities.uchicago.edu/orgs/ARTFL/

Le Dictionnaire Le Littré, Les monuments historiques de la langue française sur CD-Rom, Redon