A Festschrift for

Edmund Gussmann

from His Friends and Colleagues

Edited by

Henryk Kardela

Maria Curie-Skłodowska University, Lublin

and

Bogdan Szymanek

Catholic University of Lublin

The University Press of the Catholic University of Lublin

Lublin 1996

Distributed Morphology and the Inflection of Words[1]

Morris Halle

Massachusetts Institute of Technology, Cambridge, Mass.

In syntactic discussions we often speak as though words were the elements that constitute the terminal strings which are operated on by the syntax. It hardly needs saying that this is at best a crude approximation. It is well known, for example, that the Tense of verbs and Number and Case of nouns have the status of independent syntactic entities that require separate nodes in the terminal string, yet in many languages these morphemes merge into a single word with the adjacent stem. Moreover, when we examine the treatment of the inflections in different languages it becomes clear that the treatments are quite heterogeneous differing substantially from language to language. Languages differ, for example, as to whether or not they fuse Number and Case into a single affix. Languages also differ in whether or not they introduce into the word, affixes that have no status in either syntax or semantics. These and many similar facts lead us to conclude that the structure of the words of a language is to a significant extent independent of its syntax/semantics, and it is this fact that has traditionally justified the existence of an autonomous module in the grammar, labelled *morphology*.

The effects of the morphology module are especially obvious in languages with rich inflectional systems, where unaffixed words of the kind that are commonly encountered in English are rare or even non-existent. In some of the languages with rich inflectional systems, the different affixes correspond directly to entities such as Number and Case that play a role in the syntax and semantics of the language. Thus, as

[1] Subject to the usual disclaimers I thank S. Bromberger, A. Calabrese, C. Jakubowicz, M. Kenstowicz, A. Marantz, R. Noyer and J.-R. Vergnaud for reading and commenting on earlier versions of this paper. I also thank the audiences at MIT, the University of Venice, and at the Laboratoire de Psychologie Expérimentale in Paris, where this paper was presented in the spring of 1994.

illustrated in (1a) the Finnish word *tehta-i-ssa* 'in the factories' consists of the stem *tehta*, the Plural marker *i* and the Inessive Case marker *ssa*. Such a transparent relationship between affixes and specific syntactic/ semantic entities is typical of an *agglutinating* language like Finnish. But not all languages are of this kind. In the so-called *inflecting* languages, of which the Indo-European languages with rich inflection systems are prime examples, we find that the morphemes which compose the word do not stand in a one:one relationship with the elements that figure in the syntax and semantics. What is especially striking about such a language is that we find among its affixes some that are without correspondent in the syntax/ semantics. Both of these properties are illustrated in (1b) with forms from the Latin declension. It is readily seen that the words in (1b) are composed of a Stem, a Theme which has no correspondent in the syntax/semantics, and a terminal suffix that simultaneously expresses both Number and Case.

(1) a.

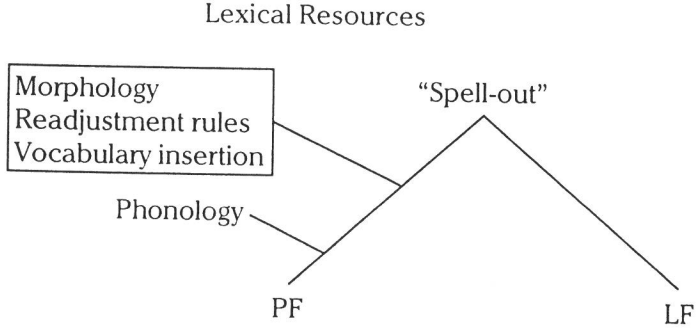
Stem	Number	Case			
tehta	i	ssa	'factory'	Pl	Inessive
kirjo	i	ssa	'book'	Pl	Inessive

b.

Stem	Theme	Number-Case	Stem	Theme	Number-Case
port	a:	s (Pl-Acc)	port	a	m (Sg-Acc)
ami:c	o:	s	ami:c	u	m
di	e:	s	di	e	m
fruct	u:	s	fruct	u	m

In (2) I have sketched the organization of the grammar that underlies this paper.

(2)



The organization in (2) reflects the basic conceptions of Distributed Morphology, the theoretical framework on which Jim Harris, Alec Marantz, Rolf Noyer, I and some others have been working in the course of the last few years. In conformity with the current version of the Minimalist Program there is no specific level of S-Structure in (2). There is, however, a point in the computation of grammatical representations of sentences where the derivation splits into two branches, of which one heads toward LF and the other toward PF. This point labelled "Spell-out" in (2) separates operations that may have an effect on the pronunciation of a sentence from those that don't. In particular, as has often been noted, movements that occur in the LF branch have no effect on the pronunciation of sentences. Implicit in this split is the further assumption that the operations that occur in the branch between Spell-out and PF are not necessarily of the same kind as those that figure in the branch headed to LF (Marantz (1994)).

While movements and other operations that occur in derivations before Spell-out affect both LF and PF, phonological properties other than those directly deriving from movements play a role only in the branch heading towards PF. In particular, phonological properties of individual stems or affixes — e.g., that *cat* begins with the phoneme /k/ and ends with /t/ — have no role in derivations prior to Spell-out, nor do they function in the operations in the right branch heading toward LF. To account for this observation, it has been postulated in DM that phonological features are supplied to morphemes in the PF branch of the derivation. Prior to that point, labelled "Vocabulary Insertion" in (2), the terminal elements of the syntactic trees that are operated on in derivations are assumed to be devoid of phonological features.

In postulating late insertion of phonological features DM deviates from other implementations of the Minimalist Program. I would like to propose here that late insertion is a special consequence of the economy principles of the Program. Specifically, we interpret these principles as limiting to a bare minimum the information that is available at each point in the computation of the sentence. Since phonological features are needed only after spell-out, the economy principles see to it that these features do not appear at an earlier point and needlessly clutter up the working space with useless information.

There is also empirical evidence for late insertion. Consider the English Past tense form *went*. This form is composed of the Past tense

suffix *t*, which occurs also in *sent, lent, bought, left,* etc., and the stem *wen/(wend)*. Both stem and suffix appear only when Infl dominates Past and only when the main verb is moved up to Infl after Spell-out. Since Past is a "weak" feature in English, the main verb does not appear next to Infl at Spell-out. In fact, because of *Procrastinate* the stem moves up to Infl only after Spell-out. Hence both parts of the form *wen-t* can only be computed in the PF branch, as implied by "late insertion."

In order to judge a theoretical proposal like DM it is necessary to have some understanding of how the proposal works technically. In what follows I have attempted to illustrate how DM treats the noun inflection in a typical IE language, i.e., Latvian, a language of which I have been a speaker since early childhood.[2] Toward the end of the paper I compare the Latvian declension with that of Latin, and I attempt to show there that in spite of the enormous genetic distance that separates Latin from Latvian, the two languages share a number of striking features which are naturally explained as the result of the preservation of specific traits of the Indo-European proto-language.

I have given in (3) the total set of forms that constitute the declension of Latvian nouns.[3]

(3) Masculine

	I	II	III	IV
	'horse'	'swan'	'stone'	'market'
Nom. Sg.	zirg-s	gulb-i-s	akmen-s	tirg-u-s
Gen.	zirg-a	gulb-j-a	akmen-s	tirg-u-s

[2] This paper supersedes earlier attempts of mine to describe the Latvian declension. See especially Halle (1992), where I have dealt with the rules of the phonology in greater detail than here.

[3] Latvian grammars recognize an Instrumental case and note that this case is identical with the Accusative in the Singular, and with the Dative in the Plural. I assume that this instance of case syncretism is due to a readjustment rule that has the effect of replacing the Instrumental with the Dative and the Accusative respectively. I believe that this can be achieved by means of Impoverishment (see below), although the details remain to be worked out. Since the rule in question applies before phonetic features are supplied to the Number-Case suffixes the Instrumental forms play no role in the discussion below and have been omitted in (3).

Dat.	zirg-a-m	gulb-i-m	akmen-i-m	tirg-u-m
Acc.	zirg-u	gulb-i	akmen-i	tirg-u
Loc.	zirg-a:	gulb-i:	akmen-i:	tirg-u:
Nom. Pl.	zirg-i	gulb-j-i	akmen-j-i	tirg-i
Gen.	zirg-u	gulb-j-u	akmen-j-u	tirg-u
Dat.	zirg-ie-m	gulb-j-ie-m	akmen-j-ie-m	tirg-ie-m
Acc.	zirg-u-s	gulb-j-u-s	akmen-j-u-s	tirg-u-s
Loc.	zirg-ua-s	gulbj-ua-s	akmen-j-ua-s	tirg-ua-s

Feminine

	I	II	III	IV
	'sister'	'earth'	'cow'	'handmill'
Nom. Sg.	ma:s-a	zem-e	guav-s	
Gen.	ma:s-a-s	zem-e-s	guav-s	
Dat.	ma:s-a-j	zem-e-j	guav-i-j	
Acc.	ma:s-u	zem-i	guav-i	
Loc.	ma:s-a:	zem-e:	guav-i:	
Nom. Pl.	ma:s-a-s	zem-e-s	guav-i-s	dzirn-u-s
Gen.	ma:s-u	zem-j-u	guav-j-u	dzirn-u
Dat.	ma:s-a:-m	zem-e:m	guav-i:-m	dzirn-u:-m
Acc.	ma:s-a-s	zem-e-s	guav-i-s	dzirn-u-s
Loc.	ma:s-a:-s	zem-e:-s	guav-i:-s	dzirn-u:-s

It is readily seen that many of the forms in (3) have the same tripartite structure as the Latin nouns in (1b). In (3) I have followed the traditional separation of the declension of Masculine nouns from that of Feminine nouns. Grammatical gender is an important factor determining the Number-Case suffixes taken by a noun. Gender, however, plays no role in the choice of Theme vowel, any of the four vowels of Latvian — /a e i u/ — can figure as Theme of nouns of both genders. The preceding is illustrated by the forms of the Dative case, which I have reproduced in (4).[4]

[4] We explain below why the Theme vowel /e/ does not surface in the Masculine declension (see rule (27)) as well as the other deviations from the canonical pattern in the Pl-Dat (see (23)).

(4)						
a.	F-Sg	ma:s-a-j	zem-e-j	guav-i-j		
	F-Pl	ma:s-a:-m	zem-e:-m	guav-i:-m	dzirn-u:-m	
b.	M-Sg	zirg-a-m	gulb-i-m	akmen-i-m	tirg-u-m	
	M-Pl	zirg-ie-m	gulb-j-ie-m	akmen-j-ie-m	tirg-ie-m	

In light of what has just been said, it is clear that in order to generate the correct inflected forms of nouns, speakers of Latvian must know not only the string of phonemes that represents the stem, but also the gender of the noun and its declension class. I have indicated the declension class information in (3) on the top line above the glosses for the masculine nouns. Without the latter "morphological" information, it is impossible to generate the different inflected forms in (3).

In spite of its crucial importance for the computation of the correct inflected forms of the noun, the Theme plays no role in either syntax or semantics. In conformity with the economy principles alluded to above, it is therefore assumed in DM that, like the phonological features, the Theme node is inserted into the terminal string only in the Morphology module. Specifically I assume that a special *Readjustment* rule modifies the syntactic structure of nouns (as well as of adjectives and verbs) by introducing a Theme constituent as an adjunct on the right of the stem. The Theme node that has thus been inserted obtains its phonetic features in the same manner as all other terminal nodes, i.e., by referring to the Vocabulary, which in DM is the repository of the phonetic information of all morphemes of the language. The phonetic information is stored in the form (5).

(5)

Theme

/a/	↔	[_____] in env.	[Class I]	+	_____
/e/	↔	[_____] in env.	[Class II]	+	_____
/i/	↔	[_____] in env.	[Class III]	+	_____
/u/	↔	[_____] in env.	[Class IV]	+	_____

The double headed arrows in (5) separate two distinct types of information: phonological features figure on the left of the arrow, whereas on the arrows' right the information is exclusively non-phonological, i.e., syntactic and morphological.

The format of (5) reflects also the fact that these entries are available for insertion into the Theme node. This is a general property of the organization of the Vocabulary in DM: the Vocabulary is subdivided into sets composed of items that can be inserted into a given terminal node (= morpheme). Interesting further properties of these "substitution" sets will be noted in the discussion of the items that are inserted into the Number-Case node.

The process of Vocabulary insertion can be pictured as an operation that copies the phonological and other features on the left of the double headed-arrow in (5) into the Theme node adjoined to the Noun stem in the Morphology module. I have illustrated in (6) the effects of Vocabulary insertion into the Theme node of two Latvian nouns.

(6)

As illustrated in (6), in Latvian the Theme is followed by an additonal suffix, which signals both the Number and Case of the noun. Latvian differs in this respect from a language such as Finnish, where, as was illustrated in (1a), Number and Case figure as separate suffixes. Since in the Latvian noun a single suffix stands for both Number and Case, we must assume that the Morphology contains a Readjustment rule that fuses Number and Case into a single terminal node into which the phonetic features of the appropriate Vocabulary items are then inserted. The Vocabulary of Latvian must therefore include in addition to the set in (5) a set of items that are inserted into the fused Number-Case node.

It is readily seen in (4) that /m/ is inserted in the Dative, except in the Singular of Feminine nouns, where /j/ must be inserted. This distribution is typical of what in phonology has been labeled as the 'elsewhere' phenomenon,[5] and like in phonology we shall employ the formal device

[5] The label was coined by Paul Kiparsky, who used it in the title of his (1973) paper "'Elsewhere' in Phonology"; the phenomenon and its treatment were discussed in SPE (Chomsky and Halle (1968)). The concept itself has well-known antecedents in Panini's grammar of Sanskrit.

of underspecification to express the 'elsewhere' relationship among Vocabulary items. In particular, as shown in (7) we shall have two entries: one for /j/ and the other for /m/.

(7)

					[# -Case]
/j/	↔	[Sg, Dat]	in env.	[Fem] +	_____
/m/	↔	[Dat]	in env.		_____

The entry for /j/ states that /j/ is to be inserted only in the Sg-Dat and only if the noun is Feminine. The context for the entry for /m/ is not specified in comparable detail: rule (7) does not state the fact that /m/ is inserted in the Pl-Dat of Feminine, and also in the Sg-Dat of Masculine nouns and in the Pl-Dat of Masculine nouns. Rather it includes no restrictions whatever, and we rely on the conventions (8), which determine how the Insertion operation works, to insure that /m/ is not inserted in the Sg-Dat of Feminine nouns.

(8) In order for a Vocabulary item to be inserted into a node in the terminal string, the item must match only a subset of the features specified in the terminal string. Insertion does not take place, however, if the Vocabulary item contains features not present in the terminal string. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal string must be chosen.

In view of (8) the two Vocabulary items in (7) compete for insertion in the Dat node, with /j/ having precedence over /m/. Thus, given the Vocabulary entries (7) and the terminal strings in (6), /j/ will be inserted in the Sg-Fem *guavij*, whereas /m/ will be inserted "elsewhere"; i.e., in the Masc-Sg *zirgam*.

We examine next the Pl-Gen forms reproduced in (9).

(9)

Masc:	zirg-u	gulb-j-u	akmen-j-u	tirg-u
Fem:	ma:s-u	zem-j-u	gov-j-u	dzirn-u

The suffix Pl-Gen suffix is evidently /u/. We capture this formally by postulating the Vocabulary entry (10).

(10)

				[# - Case]
/u/	↔	[Pl, Gen]	in env.	_____

Returning to the forms in (9) we notice that when the Theme is /i,e/ — i.e., a [–back] vowel — it surfaces as the glide /j/, but when it is /u,a/ i.e., [+back], no Theme vowel appears. This behavior is due to the operation of a pair of *phonological* rules, which are stated somewhat informally in (11).

(11) a. $V \rightarrow /j/$ in env. $\left[\begin{array}{l}\_\_\_\_\_ \\ \text{–back}\end{array}\right] + V$

b. $V \rightarrow \emptyset$ in env. $[\_\_\_\_\_] + V$

The Deletion rule (11b) accounts in part for the fact that the underlying tripartite structure of Latvian nouns does not always figure overtly in the phonetic surface string. Deletion rules, however, are not the only source of this type of asymmetry. A different reason for the absence of overt suffixes is the existence of phonological null Vocabulary items. Examples of *null* Vocabulary items are the Fem-Sg-Nom forms illustrated in (12).

(12)

Stem-Theme-Sg, Nom					Stem-Theme-Sg, Nom				
ma:s	-	a	-	Ø	zem	-	e	-	Ø

Phonetic null is also inserted in the Sg-Loc and Sg-Acc of both Feminine and Masculine nouns. We have here an example of where distinctions crucial to the syntax cease to be available for the phonology because of the working of the morphology, illustrating that the relation between syntactic and phonological representations is indirect, mediated by the morphology.

Consider now the Sg-Loc and Sg-Acc forms. In the Sg-Loc the Theme vowel is lengthened — i.e., the syllabic structure of the Theme is changed — but its feature composition remains intact. By contrast in the Sg-Acc only the feature composition of the Theme vowel changes, but the structure of the syllable remains unchanged. In (11a) both structure and features composition of the segments were affected. Rules (13a) and (13b) apply in a number of further contexts: (13a) in the Pl-Dat and Pl-Loc of Feminine nouns, and (13b) in Masculine nouns of class II.

(13) a. V → V: in env. ____
|
[Theme]+[Sg, Loc/Pl, Loc/Pl, Dat]

b. V → [+high] in env. ____
|
[Theme]+[Sg, Acc]

I interrupt the exposition at this point in order to summarize and review what has been established. I have suggested that Latvian nouns conform to the structure in (1b). The constituents of the structure in (1b) include in addition to the Noun stem, a Theme and a Number-Case suffix. The Theme has no correspondent in the syntax, and the single Number-Case node corresponds to two distinct elements in the syntax. Readjustment rules of the Morphology module introduce the Theme and fuse the Number and Case morphemes into a single terminal node. It is into these modified terminal nodes that Vocabulary items are inserted. Since the Noun is composed of three parts the Vocabulary includes three separate lists of items: a list of stems, about which nothing further will be said here, the list of entries in (5) competing for insertion in the Theme morpheme, and a third list of entries such as those in (7) and (10), which compete for insertion in the Number-Case morpheme. The insertion of Vocabulary items into the morphemes of the terminal string is subject to condition (8). The morphemes in the terminal string are also operated on by Readjustment rules of the Morphology such as those in (13), on the one hand, and by phonological rules such as those in (11), on the other. It is these three distinct modules; the Vocabulary, the Readjustment rules,

and the rules of the Phonology that together determine the surface forms of the nouns.

It may be objected that this type of account is needlessly complex. As an alternative it may be suggested that there are no morphological rules and that instead the different phonetically distinct suffixes are stored as unanalyzed units with a list of the contexts in which each unit can appear. Such a list though large and no more than a record of the observed facts would arguably be no more complicated than the principles and rules presented here. An argument against this alternative is that it makes it in principle impossible to account for the fact that the words of one language differ from those of another not only in their phonological composition, but also in morphological structure. It is a fact that Latvian and Latin nouns have Themes and in addition also a fused Number-Case suffix, whereas Finnish nouns have no Theme but have separate suffixes for Number and Case. Finally in English, like in Finnish, there is no theme suffix, but, unlike in both Finnish and Latvian/Latin, in English words may appear without suffixes altogether. These differences are, of course, not relevant for syntax/semantics, but they are clearly an essential part of the knowledge that speakers of the respective languages need and they must therefore be reflected in the grammar. It is this fact that is formally reflected in DM by the machinery of the Morphology module.

Returning to (3) we observe that Latvian has surprisingly few Number-Case suffixes. Like in all IE languages there is extensive Case syncretism in Latvian. Thus, as indicated in (14) the Fem nouns have only three distinct suffixes for the five cases of the Sg, and three for the Pl. Moreover, the suffix /s/ figures in both the Pl and Sg.

(14) Fem-Pl: /u/ /m/ /s/ Fem-Sg: /j/ null /s/

We know that in the Plural the exponent of the Gen is /u/ (cf. (10)) and /m/ is the exponent of the Dat (cf. (7)). That leaves /s/ as the exponent for Nom, Acc, and Loc, i.e., a set of Cases that could not possibly share. This radical type of case syncretism is typical of the "elsewhere" case, i.e., the complement of all the natural classes in which other Vocabulary entries have been inserted. Formally we reflect the fact that /s/ is the default exponent by representing it as the totally underspecified entry shown in (15).

(15)

/s/ ↔ [ ] in env. [# - Case] _____

In view of condition (8), item (15) can be inserted only where other, more fully specified items are ruled out. Thus, the Vocabulary items (7) and (10) are correctly given precedence over (15) in the Pl-Gen and Pl-Dat, respectively, but in all other Cases of the Plural, the item (15) is inserted.

Consider now the Singular Case endings in (14). The /j/ suffix of the Dative has been accounted for by (7). If we disregard the /s/ suffix for the moment, the null suffix is the default item for the Singular. It will therefore be represented by the entry (16).

(16)

null ↔ [Sg] in env. [# - Case] _____

The entry (16) will take precedence over the "default" entry (15), since (16) can be inserted only in Sg nodes, whereas (15) is not so restricted.

In (17) I have listed the Vocabulary items for Number-Case that have been developed to this point. The remaining two Vocabulary items that compete for insertion in the Number-Case node are given in (22b).

(17)

/j/	↔	[Sg, Dat]	in env.	[Fem] + _____	(8a)
/u/	↔	[Pl, Gen]	in env.	_____	(10)
/m/	↔	[Dat]	in env.	_____	(8b)
null	↔	[Sg]	in env.	_____	(16)
/s/	↔	[ ]	in env.	_____	(15)

The items in (17) will generate the correct outputs for all Feminine noun forms except for those that take /s/ in the Singular. One way of inserting /s/ in these case forms would be by adding special entries for the Sg-Gen and for the Sg-Nom of Class III stems. This move would imply in effect that the appearance of /s/ in these cases is a mere accident, no more likely than the appearance of /k/ or /o/ or /l/, and it would disregard completely the fact the /s/ is the default suffix. We therefore pursue a different tack and postulate the distinctness condition (18).

(18) In a set of affixes competing for insertion in a given node all entries must be phonologically distinct.[6]

It is clear that if (18) holds we can obtain the correct insertion only by modifying the feature composition of the nodes into which /s/ is inserted. Our problem then is how to modify the feature composition of the Sg-Gen and the Sg-Nom nodes so that the default /s/ is inserted into them. The obvious device here are Readjustment rules that delete the Number and Case features in the nodes under discussion. These rules are given in (19).

(19) a. Sg, Gen → Ø in env. [Fem] + [_____]
b. Sg, Nom → Ø in env. [Fem, III] + [_____]

Once the Number and Case features have been deleted the only Vocabulary entry in (17) that can be inserted in these forms is /s/, the default entry.

Readjustment rules such as (19), which delete morpho-syntactic features in the terminal string, have been called *Impoverishment* rules. In view of the conditions in (8), which govern the insertion of Vocabulary items, Impoverishment plays an important role in the morphology of dif-

[6] Work now in progress appears to indicate that condition (18) may only characterize the unmarked situation, rather than function as an absolute prohibition. This weakenning of (18) does not affect the discussion below; it only limits the contexts in which the Readjustment rules apply.

ferent languages. (See, for example, Harris (1994) on the role of Impoverishment in Spanish clitics; for additional discussion see Halle and Marantz (1993; 1994.))[7]

Before dealing with the rest of the Latvian declension, I digress briefly to consider the declension pattern of Latin nouns of the traditional fifth class illustrated in (20).

(20)

	Singular	Plural
Nom.	di-e:-s	di-e:-s
Gen.	di-e:-i:	di-e:-r-um
Dat.	di-e:-i:	di-e:-bu-s
Acc.	di-e-m	di-e:-s
Abl.	di-e:	di-e:-bu-s

The majority of forms in (20) clearly exhibit the tripartite structure (1b). The main exception are the Pl-Gen and Dat-Abl forms. Here a fourth constituent appears directly before the Number-Case suffix. I assume that this fourth constituent, which I propose to call the *Augment*, is a left adjunct of the Number-Case morpheme inserted by a special Readjustment rule. (Augments appear also in the Latvian declension, cf. the Masculine Plural forms in (3) and (23) below.)[8]

The augments thus out of the way, we see that in the Plural there are only two Number-Case suffixes: /um/ in Pl-Gen and /s/ elsewhere. (Note the parallel to the Latvian situation.) In the Sg there are four distinct suffixes: /m/ in the Acc, /i:/ in the Gen and Dat, null in the Abl, and /s/ in the Nom. In the light of (18) and the Latvian facts, the appropriate treatment of the /s/ suffix in the Sg is by means of an Impoverishment rule analogous to (19). The null suffix appearing in the Sg-Abl also has a correspondent in the Latvian declension and like its Latvian counterpart shall be treated here as the default Sg suffix. The Vocabulary items competing for insertion in the Number-Case morpheme of Class V nouns in Latin will then make up the list in (21).

[7] Impoverishment rules delete features that are crucial for LF and semantic interpretation. Thus, as the result of the application of the Impoverishment rules (19) PF strings in Latvian will lack Number information, which obviously is needed for the correct semantic interpretation. This should not be surprising, for as we have seen above the information available at PF may differ radically from that available at LF.

[8] Notice that /r/ is inserted in the Pl-Gen node after stems of Classes I, II, and V, i.e., after a Theme vowel that is [–high].

(21)

/um/ ↔ [Pl, Gen]
/m/ ↔ [Sg, Acc]
/i:/ ↔ [Sg, Gen/Dat]
null ↔ [Sg]
/s/ ↔ [ ]
} in env. [# -Case] ———

The parallelisms between (21) and (17) require no further comment. Inote that except for the /a/ of the Pl-Nom/Acc of neuter nouns, (21) is an exhaustive list of the Number-Case entries for Latin nouns.[9]

Returning to the Latvian declension, we next look at the Masculine nouns. The Sg-Gen and the Pl-Nom of the Masculine nouns require their own Vocabulary entries. This is illustrated in (22); the forms are given in (22a) and the Vocabulary entries in (22b).

(22) a.

zirg-i	gulb-j-i	akmen-j-i	tirg-i
zirg-a	gulb-j-a	akmen-s	tirg-u-s

b.

/i/ ↔ [Pl, Nom] in env. [Masc] + [# - Case] ———
/a/ ↔ [Sg, Gen] in env. [Masc] + [# - Case] ———

The Vocabulary items in (22b) will automatically be ordered before the items listed in (17) because they have heavier contextual restrictions than any of the items in (17).

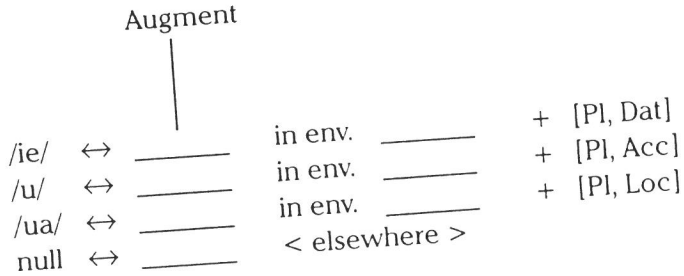
[9] The main morphological Readjustment rules for Latin are: Nom → Acc in Neuter; Impoverish Sg-Nom except in Class I; Impoverish Sg-Gen in III,IV, i.e., where Theme is [+high]; Impoverish Case only in Sg-Neut, Sg-Obl in II, III. The Theme vowel is long except in III, and phonological Readjustment rules shorten the Theme in Sg-Nom/Acc. The Theme vowel is deleted in Sg-Nom in III (except in nouns of the *civis* class), (also in Pl-Gen), after /r/ in II. Subject to the usual disclaimers, I thank A. Calabrese for providing me with most of these suggestions.

The Plural Dative, Accusative and Locative forms of Masculine nouns have an Augment intercalated between the Theme and the Number-Case suffix. The augment is seen most clearly in the forms of Class II and III.[10]

gulb-j-ie-m	gulb-j-u-s	gulb-j-ua-s
akmen-j-ie-m	akmen-j-u-s	akmen-j-ua-s

To account for the observed Number-Case augments I postulate a Readjustment rule that adjoins an Augment node as a left sister to the Number-Case node of Masculine nouns. The entries in (23) then compete for insertion into this node.

(23)



Similar augments are found in the Number-Case forms of other IE languages. As noted above in Latin augments appear in the Pl-Gen and in the Pl-Dat/Abl of certain noun classes.[11]

What remains to be accounted for is the /s/ suffix in the Sg forms of Masculine nouns. In the Masculine nouns, the /s/ suffix figures in the Sg-Nom and in the Sg-Gen of Classes III and IV. The situation being all but identical with that of /s/ suffixes in Feminine Sg noun discussed above, we account for the Masculine /s/ suffixes with the help of an Impoverishment rule similar to (19). We note that Classes III and IV have Theme vowels /i/ and /u/ respectively. In formulating the Impoverishment rule

[10] The phonological behavior of the Augment indicates that it is an adjunct of the Number-Case morpheme.

[11] In Polish and most other Slavic languages the Augment of the Pl-Dat and Pl-Inst is /m/, while that of the Pl-Loc is /x/ and that of Pl-Gen is /j/ alternating with /v/.

(24) we take advantage of the fact that these two Themes share the phonetic feature [+high].[12]

(24) [Sg, Nom] → Ø in env. $<[\text{Masc}]>_a$ + [Theme] + _____
$<[+\text{high}]>_b$

Condition: $a \vee b$

In the discussion above we have disregarded the fact that in Feminine Class III nouns the Sg-Nom suffix is /s/, rather than null. Since none of the Feminine Class IV nouns has singular forms — they are all *pluralia tantum* — we are free to speculate that Impoverishment would apply to them too. We reflect this formally by modifying (19) as shown in (25). Rule (25) states that the Number-Case node is impoverished in the SgGen of Feminine nouns and in nouns with a [–high] Theme.

(25) [Sg, Gen] → Ø in env. $<[\text{Fem}]>_a$ + [Theme] + _____
$<[+\text{high}]>_b$

Condition: $a \vee b$

The Sg-Nom forms of Masc nouns have no Theme in Class I and in Class III. Significantly, the Theme is null also in the Sg-Gen of Class III Masculine nouns, but the Theme is not null before the Plural /s/ suffix. In other words, the Theme vowel is deleted before the /s/ suffix when it represents an Impoverished Number-Case morpheme, but not elsewhere. The Theme is, however, not null before every impoverished /s/; null appears only in Masculine a-stems and in all /i/-stems, Masculine or Feminine. We express this formally by means of rule (26).

(26) $\left.\begin{array}{ll} <\text{/a/}>_a & \rightarrow \emptyset \\ \text{/i/} & \rightarrow \emptyset \end{array}\right\}$ in env. $\underline{\quad\quad}$ + /s/
[Theme] + [ ]
$<[\text{Masc}]>_b$ +

Condition: b iff a

[12] The fact that in rule (24) reference is made to both morpho-syntactic and phonological features sheds interesting light on the order in which the Readjustment rules apply. It suggests that the rules of the morphology, Vocabulary insertion as well as the Readjustment rules, apply cyclically beginning with the innermost constituent of the word.

Up to this point in the discussion we have distinguished Masculine Class II and III stems by postulating that Class II stems take as their Theme the non-high vowel /e/ whereas Class III stems take as their Theme the high vowel /i/. This has allowed us to state the rules (23), (24) and (25) in the simplest fashion possible. We must now account for the fact that in Masculine Class II nouns the surface Theme vowel is /i/ rather than /e/. Recall that rule (13b) has the effect of raising the Theme vowel. As stated rule (13b) is limited to Sg-Acc nouns. To include within its purview all Class II Masculine nouns a second context must be added as shown in(27).

(27) $V \rightarrow [+high]$ in env. $<[Masc, II]>_a + [Theme] + <Sg, Acc>_b$

Condition: $a \vee b$

It is obvious that rule (27) must apply after rule (26).

I have listed in (28) the rules and their orders that have been developed in the preceding discussion.

(28) Morphological Readjustment Rules: (19), (24), (25)
Vocabulary Insertion: (5), (17), (22b), (23)
Phonological Readjustment Rules: (13), (26) >> (27)
Phonology Proper: (11)

Concluding Remarks

It is assumed in DM that the morphemes that make up words in all languages are terminal nodes of the familiar hierarchical structures whose organization is determined by the principles and operations of the syntax. Since syntactic operations — e.g., constituent movement or the insertion of terminal nodes — may take place after Spell-out in the PF branch of the grammar, the hierarchical structure at PF may differ from that at Spell-out. Though far from insignificant, the effects of these operations are constrained by the fact that they are strictly local and must respect the hierarchical syntactic structure of the representations.

In addition to syntactic principles, principles of economy also affect the operation of the morphology. Thus, as noted above, since phonological information plays no role in the computation of LF, the economy principles of the Minimalist Program require that there be no phonological information in representations of sentences until this information is needed in order to compute PF. It is this economy principle that accounts for the late insertion of phonological features, which is one of the salient characteristics of DM.

It is further assumed in DM that the phonological features of the individual morphemes are stored in memory in the form of entries such as those illustrated in (5), (17), (22b) and (23). If this claim about the form of the Vocabulary items is correct then the rest of the grammar must operate in such a fashion as to make possible the insertion of these items into terminal nodes. Since Themes, Augments and Number-Case do not function as *single* units in the syntax, yet appear as distinct suffixes, the Morphology must generate the appropriate terminal nodes in the syntactic structure, and these operations must logically precede Vocabulary insertion.

A third assumption of DM is that the Vocabulary items that can be inserted into a given terminal node are grouped together in separate sub-lists like those in (5), (17), (22b) and (23). The Vocabulary items that are affixes are subject to the condition (18) which imposes restrictions on homophonous items that are inserted in a given node.

In accordance with the conditions in (8), a Vocabulary item that has many restrictions on the context in which it may occur is given precedence in insertion over one with fewer restrictions. Formally this is reflected by ordering the items in the Vocabulary so that in each of the sub-lists the item having the most restrictions is listed first and the one with the least restrictions is listed last. In the limit, a given item as, for example, the /s/ suffix in the noun declension of both Latvian and Latin may have no contextual restrictions. Such a maximally underspecified item represents the "default" case, and the contexts in which it is encountered do not constitute a natural class, for they are the complement of the contexts in which items listed earlier are inserted.

The underspecification of Vocabulary items is exploited by languages in highly original ways, most especially by means of Impoverishment rules, as illustrated above. Of special interest is the use of Impoverishment in order to satisfy condition (18).

Finally it was noted that an essential part of a speaker's knowledge of a language consists of knowledge of the different morphemes of the language. This knowledge includes phonological as well as syntactic, semantic and morphological features. The proposed format for Vocabulary

items is a specific proposal about the form in which this information is stored in the speaker's linguistic memory. As such it is subject to verification not only by narrowly linguistic evidence of the kind marshalled here, but also by data from studies of other aspects of human memory.

References

Bergmane, A. et al., eds. (1959) *Mūsdienu latviešu literārās valodas grammatika*, Latvijas PSRS Zinātnu Akadēmijas izdevniecība, Riga.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, The MIT Press, Cambridge, Mass.

Halle, M. 1992 "The Latvian Declension," in G. Booij and J. van Marle, eds., *Yearbook of Morphology 1991*, Kluwer Academic Publishers, Dordrecht.

Halle, M. and A. Marantz (1993) "Distributed Morphology and the Pieces of Inflection," in K. Hale and S. Keyser, eds., *The View from Building 20*, The MIT Press, Cambridge, Mass.

Halle, M. and A. Marantz (1994) "Some Key Features of Distributed Morphology," *MIT Working Papers in Linguistics* 21, 275-88.

Harris, J. (1994) "The Syntax–Phonology Mapping in Catalan and Spanish Clitics," *MIT Working Papers in Linguistics* 21, 321-353.

Kiparsky, P. (1973) "'Elsewhere' in Phonology," in S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York.

Marantz, A. (1995) "The Minimalist Program"in G. Webelhuth, ed., *Government and Binding Theory and Minimalist Program*, Blackwell, Oxford and Cambridge, Mass.

The "Passive Scale" Revisited: From Passives to Middle Voice Constructions. Evidence from English and Polish

Henryk Kardela
Maria Curie-Skłodowska University, Lublin

1. Introduction

Consider the following examples from Quirk et al. (1972, 809), illustrating the use of the Passive in English:

(1) This violin was made by my father.
(2) This conclusion is hardly justified by the results.
(3) Coal has been replaced by oil.
(4) This difficulty can be avoided in several ways.
(5) English is spoken here.
(6) We are encouraged to go on with the project.
(7) John was interested in linguistics.
(8) The modern world becomes more highly industrialized and mechanized.
(9) The house is already sold.

These examples form what Svartvik (1966) and Quirk et al. call the "passive scale". The sentences in (1) and (2) are directly related to their active equivalents:

(10) My father made this violin.
(11) The results hardly justify this conclusion.

(3) is a passive with two possible interpretations depending on whether *by oil* is treated as an "agentive" by-phrase or as an "instrumental" by-phrase, in which case *by = with*:

(12) a. Oil has replaced coal.
b. [People in many countries] have replaced coal by/with oil.