

Morphological stems: what William of Ockham really said

Mark Aronoff

Stony Brook University

While the distribution of alternants obviously requires expression in terms of the cluster of morphosyntactic properties with which they are associated, that distribution is not correlated with any 'natural class' of such properties, and cannot reasonably be argued to have any COLLECTIVE signaling function.

Martin Maiden (1992)

Zeal for the salvation of souls consumes me.

Saint Angela Merici (founder of the Ursuline Order)

1 Root and Stem

Two concepts that have figured widely in morphological analysis and theory for centuries are *root* and *stem* (or *theme*). The OED defines *root* in its linguistic sense as “an ultimate unanalysable element of language” and gives a first citation from John Palsgrave’s grammar of French printed in 1530, said to be the first French grammar printed in any language. In that same citation, Palsgrave also uses the term *theme* in much the same sense that I will use *stem* here. The OED definition of this sense of *theme* is “The inflexional base or stem of a word, consisting of the ‘root’ with modification or addition”. According to the OED, *stem* does not appear in its linguistic sense, which is given there as “the theme of a word (or of a particular group of its cases or tenses), to which the flexional suffixes are attached”, until 1851. A root is thus defined in terms of analyzability (it is unanalyzable morphologically), while a stem or theme is defined in terms of the inflectional elements that attach to it.¹ I will make the same distinction in this work.

· Thanks to Fabio Montermini and the organizers of the workshop on stems at IMM 14 in Budapest for inviting me to present this work. Thanks to Olivier Bonami for detailed comments and encouragement and to Jim Blevins and Greg Stump for valuable comments on the original version. I owe Greg Stump a special debt of gratitude for making the point about distribution and readjustment at the end of section 5.

¹ By this definition, a stem may or may not be morphologically complex, although many linguists, e.g. Matthews (1991, 64), do contrast root and stem precisely in terms of morphological complexity rather than following the definitions given here.

In recent work, David Embick and Morris Halle (Embick and Halle 2005), have questioned whether we need both concepts, root and stem, in linguistics:

Here we examine alternations that have been used in arguments that grammar must contain the ‘stem’ as a privileged object. We argue that the move to stems [from roots] is both unmotivated and problematic . . . (ibid. 1).

In this article, I will show that the morphological stem, at least in the sense given above, is helpful in understanding the morphology of Modern Standard French (henceforth French) and that the stem is implicit even in a description of French that purports to manage without it. I will also show that Embick and Halle’s attempt to do without stems is based on an incorrect understanding of the nature of both stems and the device that they invoke in order to manage without stems, the (re)adjustment rule.

It is important to underscore, as an anonymous referee reminds me, that Embick and Halle abjure stems not merely because they wish to but because a strict root-based model of morphology cannot accommodate stems (which can exist only within a lexeme-based realizational framework) and must instead try to shift the work that they do onto other devices, especially (re)adjustment rules. If, therefore, these other devices manage the work that stems do only awkwardly and at the cost of an explanation, then we have a general argument not only against (re)adjustment rules, but more importantly, against the entire root-based enterprise.

More generally, this article is meant to be a reminder that less is not always more. The following passage from William Thorburn, though almost a century old, is apt:

It is folly, to complicate research by multiplying the objects of inquiry; but we know too little of the ultimate constitution of the Universe, to assume that it cannot be far more complex than it seems, or than we have any actual reason to suppose. The value of this warning has just now received signal illustration from the very recent discovery of Chemical *Isotopes*; which has proved (*e.g.*), that what had previously been simply called "lead" is infinitely complex in its composition. This discovery ought to operate as a salutary check upon dogmatism, and the tendency to turn logical rules into ontological principles. (Thorburn 1918, 352).

2 Roots, Stems, Listemes, and Suppletion

Before proceeding further, I must be more explicit about what I mean by root and stem, expanding the OED definitions given above. Aronoff (1994) distinguishes the two within a realizational framework as follows:

[a]lthough root and stem both designate sound forms of lexemes, the most important difference between them is that a root is defined with respect to a lexeme, while a stem is always defined with respect to a realization rule. . . . A root is what is left when all morphological structure has been wrung out of a form. (ibid. 40)

Stems are more precisely defined there as follows (ibid. 39):

The **stem** of any realization rule is the form of the lexeme on/from which the rule/function performs its realization, i.e. the **phonological domain** of the rule.

While a lexeme has arbitrary properties on three dimensions – sound form, syntax, and meaning – as well as the usually arbitrary association among them, I will reserve the term *stem* for only the sound-form part of this trinity. A stem, in my use of this term, is a sound form. In particular, it is the phonological domain of a realization rule: that sound form to which a given affix is attached or upon which a given non-affixal realization rule operates.

As for root, it must be admitted that the concept does not have a major role in what Stump (2001) has termed lexeme-based inferential-realizational theories, simply because realization in such theories does not operate on roots in the sense used here. But we may translate the traditional definition into the following: the **root** of a lexeme is the default (lexical) phonological form associated with that lexeme; in what follows, I will operate with this definition, which is approximately that of both Stump (p. 33) and Matthews (1991).²

Other work on roots within lexeme-based realizational morphology has been principally concerned with root-indexed morphological and phonological phenomena. Aronoff (1976) noted that Latinate roots in English sometimes select specific variants of suffixes (e.g. *-ion* or *-tion* instead of the regular *-ation*) and sometimes exhibited peculiar variants when suffixed (e.g. *-vers* instead of *-vert* before *-ion*). Aronoff (2007) argued further that the traditional classification of Hebrew weak roots provided evidence for their analytic validity and morphological reality. In both cases, great pains were taken to demonstrate that meaning was not a factor in identifying roots. This independence of roots from meaning, while perfectly sensible within a lexeme-based framework, is not quite as congenial to root-based theories, though it should be noted that even such resolutely root-centric theories as Distributed Morphology call for roots not to be specified for lexical category (Embick and Marantz 2008). In any case, the argument that Embick and Halle make against stems does not depend on roots having meaning either.

Another important notion that is relevant to any discussion of roots and stems is the *listeme* (Di Sciullo and Williams 1987), which is any linguistic entity that must be learned, because its meaning or form is not predictable. A listeme may be as small as a single morpheme; it may be as large as an entire sentence if its meaning is not to be taken literally (like *a bird in the hand is worth two in the bush*); and it may be anything in between, like a stem, or a word, or a phrase, so long as it is somehow irregular and hence worthy of listing because it has some property that must be

² In non-lexeme-based theories of morphology, a root may also be said to have meaning. Within a lexeme-based theory, lexical meanings lie in lexemes or lexical entries and not in forms. The distinction is fine and not germane to the argument here.

learned as inhering in that entity.³ All roots are listemes, because they are arbitrary (nothing about the string *dog* predicts that it means ‘dog’), but not all stems are listemes, only those that are in some way irregular, like the past tense of English *go* or *have*. A stem that is perfectly predictable does not need to be listed. But Di Sciullo and Williams also note that the more complex a form is the less likely it is to be somehow unpredictable and hence listed, which means that a stem is more likely to be listed than an entire word or sentence. I mention listemes because any claim about stems must be separated from a claim about listemes but also because what is listed can be of diagnostic value in studying a system.

Listed stems play a major role in my analysis here. Whether the morphology of a language calls for stems has no direct connection with whether these stems must be listed but, as Andrew Carstairs-McCarthy so perceptively noted, listedness, especially as it relates to suppletion, "increasingly ... is coming to be seen as a clue to the way in which the inflected forms of a word are related to each other paradigmatically ... and are stored in the memory" (Carstairs-McCarthy 1994: 4410). So, the fact that a stem is listed and the degree of divergence of the listed stem from other forms of a lexeme may help us to understand the structure of the morphological paradigm that the stem belongs to.

Suppletion is quite generally a relation between signs, which Mel’cuk (1994) defines as follows:

For the signs X and Y to be suppletive their semantic correlation should be maximally regular, while their formal correlation is maximally irregular. (Mel’cuk 1994: 358)

The signs involved in a suppletive relation most often occupy cells in an inflectional paradigm. Boyé 2006, in fact, largely confines his discussion of suppletion to a “type of allomorphy, where some forms of a word are based on different stems that are not synchronically related.” This informal definition fits our needs in this article. Since the semantic relation between such cells is completely predictable, the difference in form where we normally find great similarity can be quite striking. Thus, one standard example of a suppletive pair of cells in a paradigm is the English verb forms *go* and *went*. Semantically, the relationship is completely regular (*went* is the past tense of the lexeme *GO*), but they share no properties of form at all. Other standard examples are the three stems of the Latin verb meaning ‘carry’, which are *fer-*, *tul-*, and *lat-* and the singular and plural of ‘man’ in Russian: [tʃelovek] and [ljudi]. Suppletion is a much more widespread phenomenon than had earlier been suspected (Veselinova 2006). Furthermore, as Mel’cuk (2000) shows, suppletion is not all-or-none, but gradable. Thus, the relation between *go* and *went* is maximally suppletive, since they share no phonological material but occupy cells in a paradigm

³ One might even wish to claim that entire conventional texts, like the United States constitution, are listemes in this sense, but the linguistic interest of such a claim is limited.

that are defined by completely regular morphosyntactic relations; the relation between *think* and *thought* is marginally less suppletive, since they share an initial segment and since the final segment of *thought*, [t], is a marker of past tense in English (albeit not a productive one; *have* and *had* are less suppletive, since, arguably, the past tense differs only in the absence of the final segment of the stem, [v]). What all suppletive inflectional relations share, though, is that something about their forms must be listed, even though their meanings are perfectly regular.

3 Paradigms, Default Inheritance, and Stem Spaces

The inflectional forms of a lexeme comprise a paradigm, which is usually thought of as a multidimensional grid whose cells are defined by the morphosyntactic features of a language (Stump 2001). The number of cells in a paradigm may vary from a theoretical minimum of one (in an uninflected language like Vietnamese) to hundreds (in a highly inflected language like Sanskrit) or even thousands, as in most analyses of Athapaskan languages. The task of morphology is to specify how each of these cells is filled. Linguists have known since classical times that the cells of a morphological paradigm are not often totally independent. Often, if we know the form of even one cell in a paradigm, we can predict the form of many if not all of the other cells. In traditional Latin grammar, the term *principal parts* is used for the forms of a verb (traditionally four in number) from which it is possible to predict all the other forms. Finkel and Stump 2007 is a recent translation and elaboration of this insight into modern terms, showing that languages may vary in how their principal parts are organized. Albright (2005, 2008) has proposed that morphological paradigms are organized around what he calls *base forms* and that, for the most part, a paradigm will have one base form.

There are various accounts of how the rest of a paradigm is filled in from its base form(s). Most make use of one type or another of *default inheritance*, a notion that linguists have borrowed from computer science. Perhaps the best example of the use of default inheritance in morphological analysis is the theory of network morphology (Corbett and Fraser 1993), which uses the computer language DATR (Evans and Gazdar 1989). The basic idea of inheritance is that individual (subordinate) members of a category inherit information from the (superordinate) category for which they are not specified individually; because a wolf is a mammal, we do not have to specify that it bears live young and nurses them, but we do have to specify all of its specifically vulpine properties. What makes *default inheritance* interesting is that an individual member may be specified in some way that contradicts its expected inheritance, in which case it will not inherit the superordinate (default) value. Another way of looking at default inheritance is to say that the superordinate fills in the values of the subordinate that are not already listed, because a listed value overrides the default. In English morphology, for example, we do not ever have to list the form of the present participle of a verb because we can always predict it from the basic form of the verb, even for the most irregular lexemes: *being* from *be*, *having* from *have*, *gyring* from *gyre*. But the same is not true of past participles, some of which have listed forms: *been*, *had*, *sung*, etc.

For verbs with listed past participles, though, it suffices to list the content of the past participle cell because this listing will override the default value for past participle: *had* prevents the inheritance of **haved*; similarly for *has*. There is a rich literature in psycholinguistics on exactly how the human mind accomplishes this default inheritance.

If we think of the forms of an inflectional paradigm of a given lexeme in a language as forming an inheritance tree, then an interesting question for inflectional morphology is what information in each node must be specified individually as opposed to the information that the node inherits from the nodes lying above it. In the simplest case, where one base form allows us to predict all others in a paradigm, we may say that only the root node, often called the *lexical representation* but which we may also think of as a morphological root in the sense defined in the last section, will be specified. In the worst case, we might have to specify completely every single cell (every terminal node of the tree). The only examples of this sort that I can think of offhand are English irregular nouns like *person/people*, where the paradigm contains only two forms, singular and plural, and both must be listed. But it is certainly true that individual cells sometimes need to be listed (think of *am, is, was, and were* in English). What about intermediate nodes in the tree between the root and the individual cells? Does it ever happen that an intermediate node in the inheritance tree is listed, in which case all its daughter nodes will inherit its listed value? One interesting observation that has emerged in the morphological literature over the last two decades is that intermediate nodes can be listed precisely when they are stems. In other words, stems may be suppletive. This is what concerns us here and is what, I believe, makes the notion of a stem most tangible in morphological theory.

4 French Verb Stems

In a series of insightful works in the last decade, Gilles Boyé and Olivier Bonami have shown that French verb paradigms are built on stems, more specifically in terms of a *stem space* that is the same for all French verbs (Boyé 2000; Bonami and Boyé 2002, 2003). The main tool that they use in their analysis is suppletion or listedness and they show that, in analyzing French verbs exhaustively, we may list either lexical representations, individual forms, or stems (as defined by the stem space), nothing else.

I would now like to show that even the analysis of verb paradigms found in a traditional French grammar book assumes that French verbs are organized around a stem space. For this purpose, I have chosen the work of the most revered of French prescriptive grammarians, Maurice Grévisse. André Gide, an icon of French literary culture, wrote the following about Grévisse's monumental *le bon usage* (in the year in which Gide won the Nobel Prize for literature):

La meilleure grammaire française, la bible de notre langue, une grammaire et une bible qui se lisent avec amusement et ravissement. *Le figaro littéraire*, 1947.

I have chosen for the basis of my discussion a more descriptive but equally comprehensive work, Grévisse and Goosse 1980.⁴ They make a strong claim about every verb having a form (*radical*) that is constant and that carries the sense of the verb and thus appear to describe French verbal morphology without recourse to stems:

Les désinences⁵ s'opposent au radical (*chant-*), qui est constant pour l'ensemble des formes d'un verbe . . . C'est le radical qui porte le sens que le verbe garde à travers toutes ses formes. (Grévisse and Goosse 227)

[the endings are opposed to the radical (*chant-*), which is constant for the entirety of the forms of a verb . . . It is the radical that bears the sense that the verb retains across all its forms.]⁶

In fact, this assertion is far from true of the description of French verb morphology that Grévisse and Goosse actually give in the grammar. Grévisse and Goosse instead assume the existence of more than one type of form of the verb lexeme (except in completely regular verbs), though they do not distinguish the forms terminologically, using only *radical* for all of them.⁷ In addition to what we would call the root, Grévisse and Goosse single out two entities that we would call stems, based on the identity of the forms that inflectional endings attach to in different cells of the paradigm, where these forms are distinct from what we may call the basic radical:

Dans le **futur simple** et dans le **conditionnel présent**, généralement on retrouve la forme de l'**infinitif**, . . . (Grévisse and Goosse 243)

[In the **simple future** and in the **present conditional**, generally we find the form of the infinitive]

⁴ Goosse was Grévisse's son-in-law and inherited the franchise. Both were Belgian, but that seems never to have been held against them in grammatical circles.

⁵ The French term *désinence* is descended directly from the Latin verb *desino* 'to cease, stop, end'. The English word *ending* is a literal translation. An ending is just that, the last inflectional element in a word, the inflectional equivalent of what has been termed *final suffix* in the literature on derivational morphology (Aronoff and Fuhrhop 2002), but we tend to forget this fact because the word has become bleached of its original meaning.

⁶ This and all other translations from French are mine.

⁷ I am indebted to Olivier Bonami and especially to Michel Roché for discussion of the terms *radical* and *thème* in the French grammatical tradition, which is quite different from the use of the terms *root*, *stem*, and *theme* in the English-language tradition. In discussing Grévisse and Goosse's analysis of French, I will accordingly use *radical* rather than *root* or *stem*. What matters most for our purposes is their recognition that a lexeme may have more than one *radical*.

Le pluriel de l'**indicatif présent**, de l'**impératif présent**, du **subjonctif présent**, ainsi que l'**indicatif imparfait** et le **participe présent** ont souvent le même radical . . . (Grévisse and Goosse 243)

[The plural of the **present indicative**, of the **present imperative**, of the **present subjunctive**, as well as the **imperfect indicative** and the **present participle** often have the same radical]

I will call the two forms that Grévisse and Goosse identify the *future form* and the *present plural/imperative plural/subjunctive plural/imperfect/present participle stem* or, for convenience, the *PIPS stem*.

4.1 The Future Form and the Future Stem

For clarity and convenience, I distinguish in this article the *future form*, which contains the marker [ʁ] that the verb endings attach to, and the *future stem*, to which this marker attaches. This terminological distinction is not meant to be standard. Let us begin with the future form, discussing its distribution, morphosyntax, and formation. Two tenses are built on the future form, the (simple) future and the (present) conditional. The endings of the conditional are identical to those of the imperfect indicative. Table 1 lists all the simple future and present conditional forms of the verb *aimer* 'love', with the future form highlighted, which consists of the verb root *aim-* [ɛm] followed by *-er* [əʁ].⁸

Future	Conditional
J' aimerai	J' aimerais
Tu aimeras	Tu aimerais
Elle aimera	Elle aimerait
Nous aimerons	Nous aimerions
Vous aimerez	Vous aimeriez
Ils aimeront	Ils aimeraient

TABLE 1

Future and conditional forms of the verb *aimer* 'love'

There are reasons to regard the conditional as the future of the past. Syntactically, there is evidence from sequence of tense. So, parallel to *Je sais qu'elle ira* 'I know that she will go', with the future, we find *je savais qu'elle irait* 'I knew that she would leave', with the conditional. Semantically, Iatridou (2000) provides a number of arguments concerning irrealis or counterfactual uses of the past tense. If we accept these claims, then we can regard the *-er-* marker as carrying the meaning 'future'

⁸ Throughout this work, I will use italics for French orthography and square brackets for surface phonological representation.

and the meanings of all forms containing this marker can be derived compositionally. Nothing in what follows, however, relies on the semantic unity of the future and the conditional.

It is a little simplistic to say that the future form always consists of the verb root followed by *-er-*, though it is true that the future form of all French verbs does end in *r* [ʁ]. The question is what does this *r* attach to? If it is indeed always the verb root, then there is no need to posit a distinct future stem. In order to explore this question further, though, we first need to know a little about the classes of French verbs.

Based solely on the phonological form of the infinitive, we can distinguish two large classes of French verbs: the class of those with infinitives ending in [-e], which is regular and productive, and which we will call the *e class* (it is traditionally called the first conjugation); and the class of those with infinitives ending in [ʁ], which is not productive and which we will call the *ʁ class*. This class may be further subdivided, based on the presence or absence of a thematic vowel *-i* [-i] and other criteria. The traditional division of French verbs follows this further subdivision, opposing verbs like *finir* with the thematic vowel *-i*, which are said to form the second conjugation, to all other *ʁ* verbs, which are said to form the third conjugation, but which in truth populate a grab-bag of irregular verbs (Bonami et al. 2008).

The vast majority of French verbs and almost all productively formed verbs fall into the *e class*. Only two (or three) verbs in this class are irregular (*aller*, (*r*)*envoyer*). We will assume that this class is the default class for French verbs and that, unless a verb is somehow specified otherwise, it will automatically fall into this class.

The only really viable other class (ca. 300 in number) is that of (second conjugation) verbs with infinitives in *-ir* that show the suffix *-iss-* [is] in the PIPS stem, e.g. *finir*, *finissons*. The class is marginally productive (*amerrir* ‘splash down’, *alunir* ‘land on the moon’), although Kilani-Schoch and Dressler (2005) claim that it is not regular.⁹ We will assume that it is regular but not productive and that its members are flagged by the thematic suffix *-i-* [i] in their lexical representation (e.g. [fin-i] as opposed to [ɛm]).

Next come verbs with an infinitive ending in *-re* [ʁ] (ca. 225 in number), which fall into a variety of subclasses, none very large. These are sometimes categorized according to the final segment of their root, e.g., *Vindre* (*craindre*, *peindre*, *joindre*). Much fewer in number (ca. 85) are the second conjugation verbs with infinitives in *-ir* that do not show the *-iss-* [is] suffix in the PIPS stem, e.g. *dormir*, *dormons* (ca. 85).¹⁰ Finally,

⁹ Bonami et al. (2008) provide psycholinguistic evidence against the simple dichotomy between the regular first conjugation and all others.

¹⁰ These can be treated formally as a subclass of verbs with the *-i-* thematic suffix whose members lack a PIPS stem. See below for further discussion.

there are verbs in *-oir* [wɑʁ] (ca. 40) like *avoir* 'have' and *vouloir* 'want'. These are quite diversely irregular.

For the (first conjugation) e class, the future form is indeed always the root plus *-er* [ɛʁ], orthographically identical to the infinitive, though not phonologically identical, since the orthographic *er* of the infinitive is phonologically [e] and not [ɛʁ]: *danser* [dɑ̃sɛ] but *danserai* [dɑ̃sɛʁ]. For these regular verbs, we do not need to list a distinct future stem.¹¹ Olivier Bonami points out to me that the schwa here appears only in e-class verbs and there are minimal pairs like *fonder/fonder*, where only the former may have a schwa, which indicates that the schwa is actually present in the future stem and optionally dropped, rather than vice versa. This shows that the future marker, which is *-r-* [ʁ], is not suffixed directly to the verb root of e class verbs, but that the future stem for all regular e-class verbs contains a schwa augment: [-ə].

For most of the members of the second and third conjugation, which we may also think of as the *ɜ* class (Plénat 1987), the infinitive and the future form are identical and consist of the verb root (plus the theme vowel *-i-* when called for) plus *-r-* [ʁ]: e.g., *finir, finirai*; *dormir, dormirai*; *souffrir, souffrirai*; *rompre, romprai*; *dire, dirai*.¹² So, for these verbs, both the infinitive marker and the future marker are *ɜ*. But within the *ɜ* class, there are a number of verbs the stem of whose future stem is irregular and hence must be listed. First, there are irregular *-ir* verbs of two types. For one type, the stem of future stem is the root and does not include the theme vowel *-i-*: *cour-ir* 'to run' *je cour-r-ai*; *cueill-ir* 'to gather' *je cueill-er-ai*; *mour-ir* 'to die' *je mour-r-ai*.¹³ Then there are irregular *-ir* verbs with a future stem that is not equal to either the root or the infinitive: *ten-ir* 'to hold' *je tiend-r-ai*; *ven-ir* 'to come' *je viend-r-ai*; *acquér-ir* 'to acquire' *j'acquér-r-ai*.¹⁴ The same is true of most *-oir* [wɑʁ] verbs: *sav-oir* 'to know' *je sau-r-ai*; *pouv-oir* 'can' *je pour-r-ai*; *fall-oir* 'be necessary that' *il*

¹¹ The only exceptions are the three irregular *-er* verbs: *aller, irai* and *(r)envoyer, (r)enverrai*, whose future stems must be listed.

¹² Assuming that that the final *-e* of the infinitive of verbs like *rompre* and *dire* is purely orthographic.

¹³ These are an instance of an exception to an exception, which, as discussed in Aronoff 1994, revert to the default. Compare emergence of the unmarked (McCarthy and Prince 1994). For these verbs, the absence of the exceptional theme vowel *-i-* only in the future causes the system to treat the verb as a member of the default e class, but only in the future. Furthermore, as Olivier Bonami points out, these verbs are subject to numerous errors: more often than not, people will say "courirai" (more marginally, "courerai") and "mourirai", which is more evidence that these verbs are more irregular than being in the second conjugation would make them.

¹⁴ At least one web site lists a regularized future form for *acquérir*: *acquérirai* (http://french.about.com/od/verb_conjugations/a/acquerir.htm), similar to the observation in the previous footnote. A Google search yields about 400 hits for the regularized form, versus about 17,000 for the standard form.

faud-r-a; *val-oir* ‘to cost, be of value’ *je vaud-r-ai*; *voul-oir* ‘to want, wish’ *je voud-r-ai*; *s’asse-oir* ‘to sit down’ *je m’assié-r-ai*. Finally, for a small number of *-oi-r* [waʁ] verbs, the future stem is equal to the root: *dev-oir* ‘must’ *je dev-r-ai*; *mouv-oir* ‘to move’ *je mouv-r-ai*; *pleuv-oir* ‘to rain’ *il pleuv-r-a*; *re-cev-oir* ‘to receive’ *je re-cev-r-ai* (also *con-cev-oir*, *dé-cev-oir*, *per-cev-oir*, *a-percev-oir*).

In summary, we can say the following: the future marker is constant in form and arguably compositional in meaning; there is no evidence that this future marker itself must ever be listed; but we must recognize an entity that the future marker *-r-* [ʁ] attaches to; and we must recognize that this entity is not always equal to the root of the verb (or the normal theme of the verb for *-ir* verbs). In all those cases where this stem is distinct from the root or normal theme, it must be lexically listed. Sometimes this stem is similar but not identical in form to the root of the verb, which might tempt one to use some sort of adjustment rule to derive one from the other, but I will show below that such a move is entirely equivalent to lexical listing. In short, when we take Grévisse and Goosse’s original observation to its logical conclusion, we see that the simple future and present conditional forms of French verbs all contain a suffix ʁ attached to a form of the verb that I have called the future stem. This stem is sometimes distinct from the root and must therefore be listed, in which case it will stand in a suppletive relation to the verb root.

The future stem is not a separate entity but rather must be listed *alongside* the default lexical representation, the root, of the verb. The lexical entry of verb *valoir* ‘be worth’ will contain both the root *val*¹⁵ and the stem *vaud* (which will be tagged as a (stem of the) future stem. Furthermore, as Boyé and Bonami emphasize in the works cited above, when we observe a verb has a (stem of the) future stem we are saying that this stem occurs only in a particular set of cells of the paradigm, not a one-off combination of cells but one that recurs across many verbs in the language. That is what we mean by the term stem-space: that the pattern of cells that contain a given stem is regularly distributed across the language.

4.2 The PIPS Stem

The other non-root verb stem that Grévisse and Goosse discuss is the *present plural/imperative plural/subjunctive (plural)/imperfect/present participle stem* or what I am calling for convenience the *PIPS stem*.¹⁶ This stem is found in all verb classes except for the regular *e* class, for which all non-future forms of regular verbs are formed directly on the verb root. Just like the (stem of the) future stem, the PIPS stem also appears, when it is listed separately, across a recurring set of

¹⁵ It will also contain at least two other stems: *vau* for the present indicative singular and *vaille* for the present subjunctive singular.

¹⁶ For some verbs, only the plural of the present subjunctive contains the PIPS stem; the singular of the present subjunctive then usually shows a separate stem that is used in the present indicative singular stem, not the root. For others, the entire present subjunctive shows the PIPS stem.

paradigmatic cells. The difference is that the distribution of the PIPS stem cannot be unified syntactically or semantically as can arguably be done for the future stem. Instead, its distribution is classically *morphomic* in the sense of Aronoff 1994, which is to say that it consists of a disjunction of cells in the paradigm. As with all morphemes, the central argument for the morphomic status of the PIPS stem is negative. There is no apparent coherent syntactic reason for tying together the cells occupied by the PIPS stem. Nor is there a semantic characterization that will cover exclusively the present indicative plural, the present active participle, the present subjunctive (sometimes only the plural), and the imperfect indicative (always both singular and plural), and so we must resort to a disjunctive enumeration of the cells that demand this stem. After all, both the present and imperfect are finite indicative tenses, but the subjunctive, although finite, is not indicative; and why does the PIPS stem occur only in the plural of the present indicative? For that matter, why does the subjunctive vary, including singular for some verbs but not for others? And finally, the present participle is non-finite.

For ‘semi-regular’ *-ir* verbs and many other verbs, the stem of plural present indicative forms (*finiss-ons*, *finiss-ez*, *finiss-ent*) is identical to the stem of all imperfect verb forms (*finiss-ais* etc.), all forms of the subjunctive, and the present (active) participle (*finiss-ant*). For these verbs, the PIPS stem is formed on the thematic *-i-* stem by means of *-ss-* [s] preceding the verb endings, so that the PIPS stem of these *-ir* verbs will consist of the verb root followed by what is sometimes called the stem augment *-iss-*. There is no need to list the *-iss-* augmented stem for ‘semi-regular’ *-ir* verbs. Instead, there is a general prediction that, if a verb has the theme vowel *-i-*, then it will have a PIPS stem and the PIPS stem will consist of the root, followed by the theme vowel, followed by *-ss-*. Historically, this stem augment originated in the Latin inchoative suffix *-sc-* (e.g. *amo* ‘I love’ vs. *amasco* ‘I begin to love, fall in love’) but has lost its inchoative meaning in all Romance languages, with various quite divergent fates in different languages (Maiden 2011). The *-iss-* augmented form constitutes the most striking piece of evidence for the PIPS stem, since it occurs on hundreds of French verbs with exactly the distribution in the cells of the paradigm described above. Forms containing a sample PIPS stem of this *-iss-* type are given in Table 2. The present indicative singular forms are highlighted, since they are not formed on the PIPS stem.

Present	Imperfect	Participle	Subjunctive
Je finis	Je finissais	finissant	Je finisse
Tu finis	Tu finissais		Tu finisses
Elle finit	Elle finissait		Elle finisse
Nous finissons	Nous finissions		Nous finissions
Vous finissez	Vous finissiez		Vous finissiez

Ils <i>finissent</i>	Ils <i>finissaient</i>		Ils <i>finissent</i>
----------------------	------------------------	--	----------------------

TABLE 2

PIPS stem forms of the verb *finir* ‘finish’

Under this analysis, we may say that those *-ir* verbs like *dormir* ‘sleep’ that do not show the augment are exceptionally specified as lacking the PIPS stem that is otherwise found in *-ir* verbs. As a result, the cells that are normally formed on the PIPS stem will instead use the verb root where the PIPS stem is called for, since that is the form available. These verbs also lack the *-i-* theme vowel and in present indicative singular forms they lack the last consonant of the root. A sample verb of this type is given in Table 3.

Present	Imperfect	(Present) Participle	Subjunctive
Je dors	Je <i>dormais</i>	<i>Dormant</i>	Je <i>dorme</i>
Tu dors	Tu <i>dormais</i>		Tu <i>dormes</i>
Elle dort	Elle <i>dormait</i>		Elle <i>dorme</i>
Nous <i>dormons</i>	Nous <i>dormions</i>		Nous <i>dormions</i>
Vous <i>dormez</i>	Vous <i>dormiez</i>		Vous <i>dormiez</i>
Ils <i>dorment</i>	Ils <i>dormaient</i>		Ils <i>dorment</i>

TABLE 3

A partial paradigm for exceptional *-ir* verbs

For many irregular verbs (those without the *-i-* theme vowel), the PIPS stem cannot be derived from the root, as it can with semi-regular *-i-* verbs, and hence must be listed as a distinct stem within the lexical entry: *craind-re, craign-ons; connaît-re, connaiss-ons; écri-re, écriv-ons; moud-re, moul-ons; résoud-re, résolv-ons*. For other irregular verbs, the PIPS stem may sometimes be equal to the root of the verb. Examples include *pouv-oir, pouv-ons; attend-re, attend-ons; cueill-ir, cueill-ons; mett-re, mett-ons; ten-ir, ten-ons*. Many verbs with listed PIPS stems also have a distinct listed present indicative singular stem (e.g., *je peux, je mets, je tiens, je crains, je meus*); some also have a listed future stem, as noted above (e.g. *je pourrai*).

Bonami and Boyé (2002) provide evidence for a number of further stem distinctions within the space of the PIPS stem, but the two most widespread stems in French are the future stem and the PIPS stem, with the latter, at least, having a purely morphomic distribution. For many verbs, there is no need to list either of these stems, since they may be predicted via a regular rule, but it is important to keep in mind that, although looking at listed forms is a good diagnostic procedure

for analyzing the stem space of a language, listedness is not a necessary criterion for stemhood.

Why Grévisse and Goosse should reject the idea that the endings of a verb may attach to more than one basic form is puzzling, since they have the term *thème* at their disposal to designate precisely the type of entity that they refer to in their discussion of what I am calling the future and PIPS stems. I suspect that their reasons are ontological: why multiply entities? It is simpler for a verb to have only one root (*radical*) rather than a multiplicity of stems (*thèmes*). We will now turn to a more explicit exposition of this ontological point of view.

5 Stems and readjustment rules

We will now return to a fuller discussion of the arguments of Embick and Halle 2005, who abjure all uses of stems, regardless of whether the stems are listed lexically or not. Instead of stems, they have recourse to readjustment rules operating on roots. Embick and Halle define readjustment rules as follows:

Readjustment rules are phonological rules that effect changes in a given morphosyntactic context and that typically include lists of Roots that undergo or trigger these changes. (6)

As an example of a readjustment rule, Embick and Halle give the following (7):

/I/ > /æ/ /X__Y [past], X = √SING, √SIT, √RING, √SINK, √BEGIN . . .

They observe further on the same page that “Readjustment rules are phonological rules: their distinguishing property is that they are conditioned by both morphosyntactic and Root-specific information.”

Aronoff 1976 gives a similar definition of what is there called *adjustment rules*:

a small class of those [phonological rules] which were previously termed morphophonemic, namely those which are restricted to specific morphemes and take place in the environment of specific morphemes. (ibid. 87)

Henceforth, I will use the acronym (re)adjustment rule as a cover term for both notions, which seem to be identical. In their article, Embick and Halle try to argue that stems are superfluous, or in their words: “that the move to stems [from roots MA] is both unmotivated and problematic . . . (ibid. 1).” I will now show that the invocation of (re)adjustment rules is, in fact, a false canard, a red herring. First, whether a grammar calls for stems has nothing at all to do with (re)adjustment rules, because (re)adjustment rules are not listed items but rather a mechanism for representing the irregular *relation* between two listed items; they do not represent the forms themselves. Second and more importantly, not all stems are listed. Thus, regardless of whether (re)adjustment rules are needed, they cannot replace either listed items or stems.

Embick and Halle work within a theory, Distributed Morphology, which makes a fundamental distinction between what they call abstract morphemes and roots. Abstract morphemes are:

... composed exclusively of nonphonetic features, such as [past] or [pl], or features that make up the determiner node D of the English definite article *the*. (ibid. 2)

Roots, by contrast:

... make up the open-class vocabulary. They include items such as $\sqrt{\text{CAT}}$, $\sqrt{\text{OX}}$, or $\sqrt{\text{SIT}}$, which are sequences of complexes of phonetic features, along with abstract indices (to distinguish homophones) and other diacritics (e.g. class features). (ibid.)¹⁷

DM is a derivational theory of morphology and syntax, in which abstract morphemes are spelled out by means of a P[honological] F[orm] derivation that combines roots and abstract morphemes and “adds phonological material to the abstract morphemes in a process called Vocabulary Insertion. During Vocabulary Insertion, individual Vocabulary Items – rules that pair a phonological exponent with a morphosyntactic context – are consulted . . .” (ibid.) Importantly, “Roots are not subject to insertion. A consequence of this view is that it is not possible for Roots to show suppletion.” (ibid. 5-6) We will return to suppletion below. The problem for the moment is that, if suppletion (listing of irregularly related forms of a lexeme or root) is forbidden, then the theory “requires a number of further rules to alter the phonology of the Root, as in the case of the past tense form sang- \emptyset . Such rules are called Readjustment Rules”. (ibid.) Later on in the article, Embick and Halle suggest that “The distinction between abstract morphemes and Roots is further marked by the important fact that the phonetic exponents of morphemes may be subject to phonological Readjustment Rules.” (ibid. 16) This statement appears to directly contradict the passage just quoted about (re)adjustment rules and roots; further, it is not clear why they make it. One likely interpretation of the sentence is that (re)adjustment rules apply only to *concatenations* of abstract morphemes and roots. A root by itself cannot undergo a (re)adjustment rule, only a root plus an abstract morpheme, or maybe two abstract morphemes in a row. This restriction (that (re)adjustment rules are always context-sensitive) is implicit in the definitions of (re)adjustment rule that both Embick and Halle and I use in the passages quoted above.

Armed with (re)adjustment rules, Embick and Halle are ready to show that stems “have no place in Distributed Morphology.” (ibid. 17) In fact, what they want to argue against is not stems but what Anderson (1992) calls *lexical stem sets*, the idea that a single lexeme or root may possess a lexically specified “group of phonologically distinct stems” (ibid. 133). Instead of a lexical stem set, claim Embick and Halle, a root or lexeme must have only one sound form; what appear to

¹⁷ Roots, in this theory, are also meaningful.

be distinct listed stems must then be generated by (re)adjustment rules. This, then, is how they wish to replace stems with (re)adjustment rules. Exactly why Embick and Halle adopt the claim that a root cannot have more than one basic form is a mystery that I will return to in the conclusion. Any analysis of their claim is made more difficult, though, by the fact that they do not make a clear distinction in their discussion between the form of a root and its status as a linguistic sign.

5.1 The power of (re)adjustment rules and of suppletion

What is the distinction between the two solution types, the one that uses stem sets and the one that uses (re)adjustment rules to generate non-basic stems? The answer, according to Embick and Halle, lies in the distinction between suppletion and phonology:

Our objection is not that apparent instances of suppletion should simply be ignored. Rather, the point is that the theory must sharply distinguish between the extremes of predictability and suppletion if it is to be explanatory in any meaningful sense. Introducing stems into the theory makes this distinction impossible to draw, or requires a host of additional stipulations. An approach that generalizes suppletion makes the weakest possible predictions concerning sound/meaning relationships: it makes no prediction. (ibid 18-19)

Although one can sympathize with the fear of generalizing suppletion that is at the heart of this passage, this fear seems to have engendered confusion between the notions of stem and listeme, which continues throughout the article. In point of fact, Embick and Halle never provide any arguments against stems, in the sense of the phonological domain of a realization rule; their arguments are only about listemes, or at least listed stems, insofar as listing is related to suppletion. Even so, their fear is misplaced, because of an erroneous claim about the nature and relative power of (re)adjustment rules and lexical listing that is implicit in the passage just quoted.

Embick and Halle evince such abhorrence of suppletion that they make the empirical claim that “It is not possible for Roots to show suppletion.” (ibid. 5-6) but that suppletion is restricted to a small semantically circumscribed set of items:

It has been observed that suppletive verbs (as well as adjectives and nouns) constitute but a vanishingly small portion of the total vocabulary of a language, and that semantically these verbs express very elementary notions. Suppletive verbs in the languages of the world have meanings such as ‘be’, ‘go’, ‘bear’, etc., but not ‘grind’, ‘withstand’, ‘animadvert’. We take it that these cases involve light-verbs, i.e. members of the functional vocabulary. (ibid. 11)

This is a strong empirical claim and it is false, as revealed by a quick glance at the suppletion database of the Surrey Morphology Group, which is conveniently available on-line at www.smg.surrey.ac.uk/Suppletion/index.aspx. The database covers only 34 languages, which makes the results all the more surprising.

There are 104 suppletive nouns listed in the Surrey database, containing the following senses:

knife, man, woman, bread fruit tree, younger brother to a man, wife, father, corner of a sack, pier of a bridge, child, shepherd, cow, car, puncture, tyre, person, house, tortilla, tree, dog, husband, animal, year, wife's father, wife's mother, elder brother's wife, son, son's wife, daughter's husband, marriageable relative, daughter of male ego's sister, son of male ego's sister, grandfather, sago palm

There are 151 suppletive verbs in the Surrey database, containing the following senses:

go, come, eat, say, become, be, give, be somewhere/have, be seated, be heavy, be (animate), see, drink, do, intend, impale, speak, be somebody, hear, take, put lying, catch, sit down, lie

There are 21 adjectives, with the following senses:

good, many, few, bad, big

Although some of these senses may be “members of the functional vocabulary,” most are not. Israeli Hebrew, one of the languages in the database, contains a suppletive verb for ‘have, exist’ (positive *yef*, negative *en*), which falls within Embick and Halle’s claim, but it also contains suppletive nouns like *tayer* ‘tire’ pl. *tsmigim*, and *oto* ‘car’ pl. *mexoniyot*. The relation between the two verbs meaning ‘say’ (*amar* and *higid*) is lovingly detailed by Myhill (2001), who concludes that they are entirely suppletive for younger speakers, with the former used exclusively in the past and present and the latter used in the future, infinitive, and imperative. Embick and Halle’s claim is obviously false, which makes one wonder why they put it on the table. I will return to that issue in my conclusion.

Embick and Halle repeatedly underscore the absence of explanatory value in suppletion. This observation is true, though it should not lead us to conclude that suppletion does not exist, only that it is undesirable at some metatheoretical level. If young speakers of Israeli Hebrew are moving towards a suppletive paradigm for ‘say’, we can’t stop them by telling them that the move is metatheoretically bad. They have not read Ockham and they do not care about metatheory or epistemological methodology.

By contrast, (re)adjustment rules, say Embick and Halle, do not come with the metatheoretical baggage that suppletion carries:

These rules have the limited expressive power of phonological rules. . . .
Readjustment Rules can therefore not be employed to relate phonetic exponents of radically different shapes. (ibid. 16)

For Embick and Halle, then, (re)adjustment rules are explanatory in a way that suppletion is not, which may be one reason why Embick and Halle would like to limit the scope of suppletion to “light-verbs” and such-like sports of nature. But is it

true that phonological rules have limited expressive power, that they cannot be employed to relate phonetic exponents of radically different shapes and are thus more explanatory than listing? The answer, which has been known for almost forty years, is no. As Johnson (1972) proved, phonological rules have the generative power of a regular relation or finite-state transducer. Using a phonological rule, we may rewrite any string of segments as any other string of segments in a specified context:

(1) *Be* > *am* /_ [1 sg pres]

Be > *was* /_ [sg past]

Epiphany > *lunchmeat* / _*dogbreath*

In short, (re)adjustment rules, and any phonological rule, have precisely the same expressive power as suppletion. There is thus no metatheoretical or 'explanatory' reason to prefer (re)adjustment rules over listing. One might imagine some evaluation metric that operated over the relations between two strings and computed the distance between them, such that the shorter the distance the more highly valued the relation. Such a metric would operationalize Mel'cuk's idea that suppletion is gradable, but this possibility only underscores the continuous connection between suppletion and regular phonological relations. They are poles of a single dimension, not devices of different power. It is therefore false to claim that (re)adjustment rules are methodologically superior to listing on the basis of mathematical power.

5.2 (re)adjustment rules are relations

In any case, the entire enterprise is misguided. (re)adjustment rules cannot do away with listed stems because (re)adjustment rules and listed stems are two entirely different animals. (re)adjustment rules and suppletion are both ways of describing the relations between forms. Listed stems are not relations. They are the forms themselves.

Everyone agrees that (re)adjustment rules are phonological rules that apply to particular roots in particular morphosyntactic contexts. This means that, for every application of a (re)adjustment rule we must specify both a root or lexeme and a morphosyntactic context. Looked at another way, every application of a (re)adjustment rule takes place in a particular paradigmatic cell. What (re)adjustment rules do is to permit us to underspecify the form that occupies the cell. (re)adjustment rules are thus a type of data compression device, a way of reducing redundancy in data, but they do not reduce *information*; each cell that undergoes a given (re)adjustment rule must be marked in some way and that means that each such cell must be on a list.

Consider how we would apply (re)adjustment rules to the three English verbs in Table 4, all of which have the basic lexical phonological form /riŋ/, but which differ in the forms of their past tense and past participle:

Spelling and sense of verb	Lexical form	Past tense form	Past participle form
<i>Ring</i> 'encircle'	riŋ	riŋd <i>ringed</i>	riŋd <i>ringed</i>
<i>Ring</i> 'sound clearly'	riŋ	ræŋ <i>rang</i>	rʌŋ <i>rung</i>
<i>Wring</i> 'compress by twisting'	riŋ	rʌŋ <i>rung</i>	rʌŋ <i>rung</i>

Table 4: Some English verb forms

The first *ring* is completely regular: all we need to list in order to obtain the three forms in question is its default lexical form. The second *ring* belongs to the *sing* class of irregular verbs and so must be marked to undergo two separate (re)adjustment rules, one for each of the cells in question. The third verb undergoes the same (re)adjustment rule in both cells, like *fling*. In the language of defaults, we would say that all the cells of the first sense of *ring* are devoid of information except for one, and that these empty cells are filled in with default forms, while the past tense and past participle cells of both the other roots contain information that overrides the default forms. What really matters when we speak of listing is that a cell contains information that overrides the defaults, not exactly what this information is or how it is encoded, whether it consists of a complete form or of a pointer to a (re)adjustment rule.

I hope it is clear by now that (re)adjustment rules, listing, and stems are different kinds of things entirely, though they interact. (Re)adjustment rules are relations between pairs of listed items. What relates listing and stems (as I showed at length in section 3 above) is only the fact that some stems must be listed, which in turn provides evidence for the place of stems in the inheritance structure of a paradigm. (re)adjustment rules are thus at two removes from stems: they are statements about the organization of listed cells that are sometimes occupied by stems. This is what I meant above when I said that the entire discussion of (re)adjustment rules in Embick and Halle is a red herring. Their argument is not about stems at all; it is about the best way to represent the content of listed cells within a theory that insists on a single lexical representation for all roots or lexemes. Whether stems are needed and whether (re)adjustment rules are useful are completely tangential matters. That is not to say that (re)adjustment rules are of no interest at all to morphologists. In fact, like listing itself, they may provide a heuristic window into the morphological structure of a language: the (re)adjustment rules help to identify cells in the morphological paradigm. The fact that the past tense and past participle of some verbs participate in (re)adjustment rules in English is a clue to both linguists and language learners that English verb paradigms contain past tense and past participle cells. (re)adjustment rules thus perform the same learning function as suppletion, which is not surprising, since they are the same thing. But (re)adjustment rules, like suppletion and listing, are never general. They therefore cannot encode general facts about any given stem, which are general facts about the form of the lexeme in sets of cells.

The PIPS stem, for example, although its form may vary depending on the exact verb whose PIPS stem is in question, is the stem that generally occurs in the following forms of the verb: present plural/imperative plural/subjunctive (plural)/imperfect/present participle. (Re)adjustment rules have nothing to say about this distribution, which is part of the general grammar of French verbs. As Greg Stump points out to me, if we try to encode the PIPS stem into readjustment rules in a root-based theory, then every single one of these readjustment rule will operate in exactly the same environment, the set of cells in which PIPS stem occurs, as the following sample set from Stump demonstrates:

- (2) $wa > yv / X_Y$ $\left\{ \begin{array}{l} \text{[present indicative plural]} \\ \text{[imperative plural]} \\ \text{[present subjunctive plural]} \\ \text{[imperfect indicative]} \\ \text{[present participle]}, \end{array} \right.$
 $X = \sqrt{\text{BOI}} \text{ (boire)}$
- $o > al / X_Y$ $\left\{ \begin{array}{l} \text{[present indicative plural]} \\ \text{[imperative plural]} \\ \text{[present subjunctive plural]} \\ \text{[imperfect indicative]} \\ \text{[present participle]}, \end{array} \right.$
 $X = \sqrt{\text{VAU}} \text{ (valoir)}$
- $\emptyset > uv / X_Y$ $\left\{ \begin{array}{l} \text{[present indicative plural]} \\ \text{[imperative plural]} \\ \text{[present subjunctive plural]} \\ \text{[imperfect indicative]} \\ \text{[present participle]}, \end{array} \right.$
 $X = \sqrt{\text{MEU}} \text{ (mouvoir)}$
- $wa > \text{ə}v / X_Y$ $\left\{ \begin{array}{l} \text{[present indicative plural]} \\ \text{[imperative plural]} \\ \text{[present subjunctive plural]} \\ \text{[imperfect indicative]} \\ \text{[present participle]}, \end{array} \right.$
 $X = \sqrt{\text{DOI}} \text{ (devoir)}$

The theory, unfortunately, provides no place for the generalization contained in these examples to be expressed. Indeed, it does not permit such a generalization to exist. This is a case in which Ockham's razor truly applies, and in its simplest form: without allowing stems, in this case the PIPS stem, we are forced to repeat the same information dozens of times; with the PIPS stem, we state the same information once.

To repeat Maiden's insightful observation with which I opened this article: While the distribution of alternants obviously requires expression in terms of the cluster of morphosyntactic properties with which they are associated, that distribution is not correlated with any 'natural class' of such properties, and cannot reasonably be argued to have any COLLECTIVE signaling function.

The French PIPS stem is not a theoretical notion but an empirical fact: a large number of French verbs have a special stem that occurs in just this environment, sometimes produced by rule (as in regular second-conjugation verbs) and sometimes lexically listed (as in the examples just above). As Martin Maiden has shown in many publications, special stems like this are common in Romance and they persist and change through time like any other grammatical entities. A theory that can accommodate them only through numerous accidental repetitions of the same set of environments cannot be right.

6 Conclusion

What is it about stems and listemes that so exercises E &H? The answer, I believe, is a deep-rooted desire for one-to-one form-meaning correspondence in language. P. H. Matthews portrays this desire in the following passage from his book, *Grammatical theory in the United States from Bloomfield to Chomsky*:

[Chafe's 1970] account rests on the myth of a linguistic Golden Age, in which each utterance is a sequence of perfect signs uniting one form with one meaning. In their evolved or fallen state, languages have lost this simplicity. But the system of signs survives at an underlying level, and the task of the linguist is to reconstruct a 'fictitious agglutinating analog' of the actual phonetic forms (Lounsbury 1953). (Matthews 1993, 102)

Tracing the origins of this wish for a perfect language lies beyond the scope of my work here. The same impulse clearly lies behind Grévisse and Goosse's definition of the *radical* as a form that is constant in both form and meaning, despite their own evidence to the contrary. For morphology in general, Lounsbury, on the same page from which the quoted passage is drawn, traces to the early twentieth-century work of Boas, Sapir, and their students this view of what he calls *morphophonemic forms* as "'ideal' morphemes, and true units of the language, which, however, when put together into constructions have changes wrought upon them so that they and their combinations are 'actualized' or 'realized' in an alternate form." Another source is certainly Roman Jakobson:

One of the fundamental concepts in the development of modern linguistics was the idea of invariance, first recognized by the Kazan' school at the close of the 1870's, simultaneously and in parallel with the success of the same idea in mathematics. Whereas the first stage of these inquiries gave birth to the theory

of the phoneme, i.e. of the invariant on the level of sound variations, now the urgent need has arisen to establish and explicate grammatical invariants. (Jakobson 1984/1958)

Jakobson spent his entire linguistic career in a quest to find linguistic invariants. This privileging of invariance over all else drove much of classic generative phonology and, to a lesser extent, classic generative syntax. Methodologically, such a quest is hard to argue against. What I hope to have shown here is that the quest for invariance, like Saint Angela's zeal for the salvation of souls, has its pitfalls.

6.1 Ockham's Razor

William of Ockham (c. 1288 – 1348) was an English-born philosopher known best to the modern world for the methodological principle known as Ockham's Razor. The usual formulation of this principle is *entia non sunt multiplicanda praeter necessitatem* (entities are not to be multiplied beyond necessity) but Ockham never wrote these actual words. The closest Ockham came (Thorburn 1918) was *numquam ponenda est pluralitas sine necessitate* ([a] plurality must never be posited without need) or (my favorite formulation) *frustra fit per plura quod potest fieri per pauciora* (it is pointless to do with more what can be done with less), either of which can be found almost identically worded in the work of John Duns Scotus (c. 1265 – 1308), Ockham's teacher, who regarded it as Aristotelian in origin. The latter version was famously repeated by Newton in the third edition of *Principia Mathematica* (1726) without attribution. I go into this detail simply to demonstrate, following Thorburn, that the principle was neither original with Ockham nor, at least before the last century, regarded as his invention, but rather as a time-honored methodological warning. Nor should it be seen as more than that. Paul Vincent Spade, in his article on Ockham, gives a similar warning to the passage from Thorburn that I cited above:

Ockham's Razor, in the senses in which it can be found in Ockham himself, never allows us to *deny* putative entities; at best it allows us to refrain from positing them in the absence of known compelling reasons for doing so. In part, this is because human beings can never be sure they know what is and what is not "beyond necessity"; the necessities are not always clear to us. (Spade 2008, 4.3)

As Spade shows, it is the latter principle that was dearest to Ockham, not his razor. Ockham was best known in his own time as a theologian, not a logician. His own great quest, like Newton's, was to understand god. Both their warnings against multiplying entities must be understood in this theological context: if we wish to know god in our fallen state of original sin, we must be careful to posit only those entities that are necessary, those that we are forced to posit. By this, neither Ockham nor Newton meant methodologically necessary, but rather empirically necessary. Parsimony never entirely determines discovery. What I hope to have shown here is that even if our goals are much more modest, if we wish to understand not god but merely language, it is necessary to acknowledge certain entities beyond those driven entirely by methodological parsimony.

REFERENCES

- Albright, Adam. 2005. The morphological basis of paradigm leveling. Laura J. Downing, T. Alan Hall, and Renate Raffelsiefen, (eds.) *Paradigms in phonological theory*. 17- 43. Oxford: Oxford University Press.
- Albright, Adam. 2008. Inflectional paradigms have bases too. In Andrew Nevins and Asaf Bachrach (eds.), *The Bases of Inflectional Identity*. 271-312. Oxford: Oxford University Press.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark. 1994. *Morphology by itself*. Cambridge, MA: MIT Press.
- Aronoff, Mark. 2007. In the beginning was the word," *Language* 83. 803-830.
- Aronoff, Mark, and Nanna Fuhrhop. 2002. Restricting suffix combinations in German and English. *Natural Language and Linguistic Theory* 20: 451-490.
- Bonami, Olivier, Gilles Boyé, Helene Giraud, and Madeleine Voga. 2008. Quels verbes sont réguliers en français? In *Actes du premier Congrès Mondial de Linguistique Française*. 1511-1523.
- Bonami, Olivier, and Gilles Boyé. 2002. Suppletion and dependency in inflectional morphology. In *The Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*. 51-70. Stanford: CSLI Publications.
- Bonami, Olivier, and Gilles Boyé. 2003. Supplétion et classes flexionnelles dans la conjugaison du français. *Langages* 152: 102-126.
- Boyé, Gilles. 2000. *Problèmes de morphophonologie verbale en français, espagnol et italien*. Thèse de l'Université Paris 7.
- Boyé, Gilles. 2006. Suppletion. In Keith Brown (ed.), *Encyclopedia of language and linguistics (2nd edition)*. Vol. ? 297-299.
- Carstairs-McCarthy, Andrew. 1994. Suppletion. In R.E. Asher and J.M.Y. Simpson (eds.), *The Encyclopedia of Language and Linguistics*. vol. 8:4410-11. Oxford: Pergamon.
- Chafe, Wallace. 1970. *Meaning and the structure of language*. Chicago: University of Chicago Press.
- Corbett, Greville G. and Norman M. Fraser. 1993. Network morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics* 29. 113-42.
- Di Sciullo, Anna Maria, and Edwin Williams. *On the definition of word*. Cambridge, MA: MIT Press.
- Embick, David, and Morris Halle. 2005. On the status of *stems* in morphological theory. In Twan Geerts, Ivo van Ginneken, and Haike Jacobs (eds.), *Romance Languages and Linguistic Theory 2003*. Amsterdam: John Benjamins.

- Embick, David and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39. 1-53.
- Evans, Roger, and Gerald Gazdar. 1996. DATR: A language for lexical knowledge representation. *Computational Linguistics*, 22. 167-216.
- Finkel, Raphael, Gregory and Stump. 2007. Principal parts and morphological typology. *Morphology* 17. 39-75.
- Grévisse, Maurice, and André Goosse. 1980. *Nouvelle grammaire française*. Paris: Duculot.
- Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31. 231-270.
- Jakobson, Roman. 1984/1958. Morphological observations on Slavic declension (the structure of Russian case forms). In Linda Waugh and Morris Halle (eds.), *Russian and Slavic grammar: studies, 1931 - 1981*. 105-132.
- Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. The Hague: Mouton.
- Kilani-Schoch, Marianne, and Wolfgang Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen: Gunter Narr Verlag.
- Lounsbury, Floyd. 1966/1953. The method of descriptive morphology. In Martin Joos (ed.), *Readings in Linguistics*. Chicago: University of Chicago Press. 379-385. (extracted from Oneida verb morphology. 1953. Yale University Publications in Anthropology no. 48)
- Maiden, Martin. 1992. Irregularity as a determinant of morphological change. *Journal of Linguistics*, 28: 285-312.
- Maiden, Martin. 2011. Morphophonological innovation. In M. Maiden, JC Smith and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*. Vol I. Structures. 216-267. Cambridge: Cambridge University Press.
- Matthews, P. H. 1991. *Morphology (2nd edition)*. Cambridge: Cambridge University Press.
- Matthews, P. H. 1993. *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.
- McCarthy, John, and Alan Prince. 1994. The emergence of the unmarked: optimality in prosodic morphology. *Proceedings of NELS* 24. 333-379.
- Mel'cuk, Igor. 1994. Suppletion: towards a logical analysis of the concept. *Studies in Language* 18. 339-410.
- Mel'cuk, Igor. 2000. Suppletion. In Geert Booij, Christian Lehmann, and Joachim Mugdan, (eds.), *Morphologie/Morphology. Ein internationales Handbuch zur Flexion und Wortbildung/An international handbook on inflection and word-formation*. 510-522, Berlin: Mouton de Gruyter.

- Myhill, John. 2001. Suppletion, lexical meaning, semantic primitives, and translation data. *Linguistics* 39: 761-802.
- Plénat, Marcel. 1987. Morphologie du passé simple et du passé composé des verbes de l' "autre" conjugaison. *ITL Review of Applied Linguistics* 77-78. 93-150.
- Spade, Paul Vincent. 2008. William of Ockham. *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/ockham/>>.
- Stump, Gregory. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Thorburn, William. M. 1918. The myth of Occam's razor. *Mind* 27(107). 345-353.
- Veselinova, Ljuba N. 2006. *Suppletion in verb paradigms: bits and pieces of the puzzle*. Amsterdam: John Benjamins.