Suppletion and dependency in inflectional morphology

Olivier Bonami
Université Rennes 2 and CNRS UMR 7110
olivier.bonami@uhb.fr

Gilles Boyé
Université Nancy 2
gilles.boye@univ-nancy2.fr

The purpose of this paper is to present a general approach to verbal inflection with special emphasis on suppletion phenomena. The approach is applied to French in this paper, but it extends straightforwardly to other languages.[1]

The first part of the paper describes an analysis of suppletion in inflectional morphology with two design requirements. First, we attempt to provide an analysis which not only accounts for the existence of suppletion phenomena, but also accounts for the fact that suppletion is not erratic: suppletive forms tend to always appear in groups, in definite areas of verbal paradigms. Second, we try to minimize the quantity of redundant phonological information that has to be listed in the lexicon for a given lexeme. We assume that an optimal analysis of inflection should be able to derive all and only predictable inflectional forms from a single representation.

Our analysis is based on the observation of a number of *dependency relations* between inflectional forms of verbs.[2] We define for each language a *stem dependency tree* based on these observations, which allows one to predict the whole paradigm of every verb in the language on the basis of a minimal number of idiosyncratic stems.

The second part of the paper attempts to integrate the analysis in an HPSG hierarchical lexicon. Morphological dependency relations are represented directly by mentioning a lexical sign in another sign's lexical entry. The approach to suppletion proposed in the first part is made explicit using a combination of online type construction and default constraints on the phonology of dependent signs.

1 Inflectional dependencies in French

In this first section, we present our analysis of French verbal inflection. Starting with a very simple-minded view of inflection (1.1), we justify the postulation of a *stem space* as part of the paradigm of a French verb (1.2). Each slot in the stem space can be occupied by a distinct stem, but the slots are not independent of one another (1.3): we observe a number of simple (1.4) or complex (1.5) dependency relations between the stem slots, which can be used to avoid redundancy in the

[1] The first part of this paper is a streamlined version of the analysis of French verbal inflection proposed in (Boyé, 2000), which also includes analyses of English, German, Italian and Spanish along the same lines.

[2] These relations are somewhat similar to the *rules of referral* in (Zwicky, 1985, 1991; Stump, 1993).

Proceedings of the 8th International HPSG Conference, Norwegian University of Science and
Technology (3-5 August 2001)
Frank van Eynde, Lars Hellan and Dorothee Beermann (editors)
2002.CSLI Publications. http://csli-publications.stanford.edu/

finite forms

	1sg	2sg	3sg	1pl	2pl	3pl
present	Ø	Ø	Ø	ɔ̃	e	ə
imperfective	ɛ	ɛ	ɛ	jɔ̃	je	ɛ
simple past	Ø	Ø	Ø	mə	tə	rə
future	re	ra	ra	rɔ̃	re	rɔ̃
subjunctive	ə	ə	ə	jɔ̃	je	ə
subj. imperfective	sə	sə	Ø	sjɔ̃	sje	sə
conditional	rɛ	rɛ	rɛ	rjɔ̃	rje	rɛ
imperative	–	Ø	–	ɔ̃	e	–

non-finite forms

infinitive	present participle	past participle
r	ɑ̃	Ø

Table 1: French inflectional suffixes

specification of the paradigm of a verb. We conclude (1.6) by considering how our analysis extends to other languages and other types of inflectional irregularity.

1.1 Two types of irregular inflection

We assume the very simple and general hypothesis on regular inflection stated in (1).

(1) **Regular inflection**
The phonology of a regular inflectional form is a function of the phonology of some stem of the verb this form instantiates.

Notice that the hypothesis is neutral with respect to the debate between phrase-structure based (Lieber, 1980; Selkirk, 1982; Spencer, 1991) and realization-based (Matthews, 1972; Zwicky, 1985; Anderson, 1992) conceptions of morphology; notice also that it is neutral as to whether a verb has one or more than one stem.

In French, the functions deriving regular inflectional forms are all functions which suffix some phonological material to the phonology of the stem. The suffixed material for all 48 French inflectional forms is listed in table 1.[3]

This simple view of regular inflection allows us to distinguish two ways for an inflectional form to be irregular: an inflectional form may be irregular in not exhibiting the effects of the phonological function associated with its slot in the paradigm; or it may be irregular in being based on an unexpected stem. The first kind of irregularity we call *inflectional form suppletion*.

[3] Ø denotes the empty string. The only case where the inflectional function is not straightforwardly concatenative is that of the infinitive, which surfaces as e with verbs in the traditional first group, and as r with all other verbs; Boyé (2000) provides a unified analysis of the infinitive explaining away the surface difference. Notice also that to improve readability, the suffixes listed in table 1 ignore latent consonants which surface only in *liaison* contexts.

Lexeme	infl. class	suppletive form	ungrammatical reg. form
être 'be'	prst. 1sg	*suis* (sɥi)	**es* (*ɛ)
	prst. 1pl	*sommes* (sɔm)	**étons* (*etɔ̃)
	prst. 2pl	*êtes* (ɛt)	**étez* (*ete)
	prst. 3pl	*sont* (sɔ̃)	**êtes* (*ɛt)
avoir 'have'	prst. 1sg	*ai* (ɛ)	**as* (*a)
	prst. 3pl	*ont* (ɔ̃)	**avent* (*av)
faire 'do'	prst. 2pl	*faites* (fɛt)	**faisez* (*fəze)
	prst. 3pl	*font* fɔ̃	**faisent* (*fɛz)
	imperative 2pl	*faites* (fɛt)	**faisez* (*fəze)
aller 'go'	prst. 1sg	*vais* (vɛ)	**vas* (*va)
	prst. 3pl	*vont* vɔ̃	**allent* (*al)
dire 'say'	prst. 2pl	*dites* (dit)	**disez* (*dize)
	imperative 2pl	*dites* (dit)	**disez* (*dize)

Table 2: French suppletive inflectional forms

Table 2 is an exhaustive list of the clear suppletive inflectional forms of French.[4]

There is clearly another type of irregular inflection, which is illustrated in (2). In the present tense, all forms of both *laver* ('to wash') and *mourir* ('to die') exhibit the effect of the phonological function associated with its slot; thus no form of these two verbs is a suppletive inflectional form. But there is a contrast between the two verbs: all forms of *laver* are based on the same stem, whereas two different stems[5] are used by *mourir*. Thus *mourir* is irregular in having multiple stems where regular verbs have only one.

(2) a. *laver* ('wash')

	sg	pl
1	lav+Ø	lav+ɔ̃
2	lav+Ø	lav+e
3	lav+Ø	lav+ə

b. *mourir* ('die')

	sg	pl
1	mœr+Ø	mur+ɔ̃
2	mœr+Ø	mur+e
3	mœr+Ø	mœr+ə

The data in (2) shows that some French verbs have multiple unrelated stems. The next question is to see how how many stems a verb may have, and which part of the paradigm a stem is used to build.

[4]The careful reader will notice that we include in table 2 the forms *suis*, *ai* and *vais*, whose status of suppletive inflectional forms can be debated: since the phonological function associated with the present-1sg is the identity function, one can never tell whether the function has applied to a form or not; see note 7.

[5]It is important to note that these two stems are not related by any productive phonological rule of French. In the remainder of this article, every time we speak of multiple stems of a single verb, we mean that the differences between these stems cannot be explained by regular phonological alternations known to be productive in contemporary French. We differ in this matter with previous works, such as Van Den Eynde and Blanche-Benveniste (1970) or Fradin (1993), who attempted to make generalizations about all phonological alternations, and Plénat (1984, 1987) who attempted to make phonological generalizations about the non-productive verb classes.

1.2 The stem space

Assuming that the inflectional forms of a single verb can be built on two or more stems, it is a logical possibility that a verb have one distinct stem corresponding to each of its inflectional forms; even if this is excluded because such a verb would be too idiosyncratic to be learned, it is a possibility that different verbs base different collections of inflectional forms on the same stem. But this is not what we observe: an exhaustive examination of French verbs shows that some parts of the paradigm are always build using a unique stem.[6] This is shown in (3) with the present tense paradigm. There are verbs which use up to three different stems in the present, but all verbs use the same stem for the three singular forms, and all verbs use the same stem for the first and second person plural.[7]

(3)

Verb	1sg	2sg	3sg	1pl	2pl	3pl
laver ('wash')	lav+Ø	lav+Ø	lav+Ø	lav+ɔ̃	lav+e	lav+ə
asseoir ('sit')	asje+Ø	asje+Ø	asje+Ø	asɛj+ɔ̃	asɛj+e	asɛj+ə
mourir ('die')	mœr+Ø	mœr+Ø	mœr+Ø	mur+ɔ̃	mur+e	mœr+ə
boire ('drink')	bwa+Ø	bwa+Ø	bwa+Ø	byv+ɔ̃	byv+e	bwav+ə

To account for these groupings, we assume that French is characterized by a *stem space* which consists of a number of slots; each slot in the inflectional form paradigm selects a slot in the stem space. For example, the data in (3) shows that the stem space for French consists of at least three slots. All present singular inflectional forms select the same slot, which we may call the *present-sg stem slot*. A distinct slot is selected by present first and second plural forms, and a third slot is selected by the present third person plural. A given verb may fill different slots with the same stem, but each pair of stem slots is justified by the fact that at least one verb places distinct stems in the two slots.

Generalizing the strategy used to account for the present tense to the whole paradigm of all French irregular verbs, we obtain a stem space with 12 slots, which is described in detail in table 3.[8]

For reasons that will be made explicit in paragraph 1.6, in the remainder of this paper, we focus on the first 8 of the 12 stem slots listed in table 3, which entails that that we will only account for 28 of the 48 distinct inflectional forms of French verbs.[9]

[6] A similar study was done in (Swiggers and Van Den Eynde, 1987) to establish a typology of irregularity in French.

[7] Notice that the verbs *être* ('be'), *avoir* ('have') and *aller* ('go') are apparent exceptions to this observation: all three verbs have different forms in the first person singular (respectively sɥi, ɛ and vɛ) and second an third person singular (respectively ɛ, a and va). However, as stated in note 4, we assume that the three incriminated forms are suppletive inflectional forms, and thus are not based on any stem of the verb.

[8] Notice that in four cases, we are led to postulate a stem slot which is used to build only one inflectional form. However, they are clearly distinct objects; for instance, the present participle *stem* of the verb *savoir* ('know') is saʃ, while the present participle *form* is saʃɑ̃.

[9] It should be noticed that the stem slots we are led to postulate do not always correspond to syntactically and/or semantically coherent classes of inflectional forms. For instance, a single stem slot accounts for all imperfective forms and just first and second person plural present forms; likewise, a single slot is responsible for all subjunctive forms except the first and second person plural. What this shows is that the stems that are involved in inflectional morphology cannot be taken to be the phonological exponence of syntactico-semantic feature bundles in the way proposed for example in Stump (1993). Rather, they are pure morphological objects in the sense of Aronoff (1994).

stem name	inflectional forms build on this stem
imperf./prst. 12pl	all forms of the imperfective; first and second person plural in the present
prst. 3pl	present third person plural
prst. sg	all singular forms in the present
subj. 12pl	first and second person plural in the subjunctive present
subj. sg/3pl	all singular forms and third person plural in the subjunctive present
imper. sg	imperative second person singular
imper. pl	imperative first and second person plural
prst. part.	present participle
fut./cond.	all forms of the future and the conditional
s.pst.	all forms of the simple past and subjunctive imperfective
inf	infinitive
pst. part.	past participle

Table 3: The French stem space

1.3 Filling the stem space

Assuming that the French stem space has 8 slots, it remains to be stated how the grammar of French fills these slots for each verb in the language. A first requirement we can put on the grammar is that each distinct stem should be listed only once. Given that all French verbs fill the 8 slots of the stem space with 1 to 3 distinct stems, [10] we should thus be able to describe the stem space of each verb with no more than three lexical entries.

An obvious way to limit the number of stems which have to be listed in the lexicon is to use lexical entries which are underspecified as to the stem slots they apply to. For instance, assuming that the stem slot an entry applies to is indicated by a feature STEM-SLOT,[11] we could assume the following two lexical entries for the two stems of the verb *asseoir* ('sit').[12]

(4) a. $\begin{bmatrix} \text{PHON} & \text{asje} \\ \text{STEM-SLOT} & \textit{prst-sg} \vee \textit{imperative-sg} \end{bmatrix}$

b. $\begin{bmatrix} \text{PHON} & \text{asɛj} \\ \text{STEM-SLOT} & \neg(\textit{prst-sg} \vee \textit{imperative-sg}) \end{bmatrix}$

The problem with such an approach is that it does not restrict the ways of filling the stem space in any way. If we simply allow lexical entries to account for any combination of slots, we predict that the number of conjugation patterns found in the language is very high: assuming that

[10] If we take into account all 12 stem slots, the observation is that no verb fills the 12 slots with more than 6 distinct stems.

[11] This hypothesis is made here for expository purposes only; in the analysis presented in section 2, stem slots correspond to types of lexical signs, and no STEM-SLOT feature is used.

[12] These descriptions could be improved upon by (i) relying on a hierarchy of STEM-SLOT values to avoid the disjunction in (4a) and (ii) relying on some encoding of the Elsewhere Condition to avoid the STEM-SLOT specification in (4b). We do not attempt such an improvement here, since this kind of lexical entry will be abandoned shortly.

example verb	imperf. prst.12pl	prst.3pl	prst.sg	subj.12pl	subj.sg subj.3pl	imper.sg	imper.pl	prst.part
laver ('wash')	A	A	A	A	A	A	A	A
devoir ('owe')	A	B	B	A	B	B	A	A
haïr ('hate')	A	A	B	A	A	B	A	A
avoir ('have')	A	A	A	B	B	B	B	B
être ('be')	A	A	B	C	C	C	C	A
valoir ('cost')	A	A	B	A	C	B	A	A
savoir ('know')	A	A	B	C	C	C	C	C
pouvoir ('can')	A	B	B	C	C	B	A	A
vouloir ('want')	A	B	B	A	C	B	A	A
faire ('do')	A	A	B	C	C	B	A	A

Table 4: The 10 French conjugation patterns

every verb has at most three distinct stems, there are 1094 distinct ways of filling the stem space that are predicted to be possible by this approach.[13] Given that French has approximately 350 irregular verbs, it cannot be expected that all 1094 conjugation patterns are actually attested; but it is expected that some reasonably large subset of the possible patterns are attested. It is thus surprising to notice that only 10 distinct ways of filling the stem space are actually attested in French. These ten ways are listed in table 4.

We take this discrepancy to indicate that there must be constraints on the way a verb fills the stem space. The following two paragraphs exhibit two kinds of *dependency relations* between stem slots that constrain the stem space.

1.4 Simple stem dependency

As first observed in (Morin, 1987), there is no verb in French which has the same stem for the imperfective and for the present singular, but has a different stem for the present third plural. This is shown in table 5.

Morin proposes that there is a a dependency relation between the three stems used in the present tense. The present third plural is either suppletive or identical with the imperfective; and the present singular is either suppletive or identical with the present third singular:[14]

(5) imperf/prst.12pl $\longrightarrow$ prst.3pl $\longrightarrow$ prst.sg

The two binary dependencies allow for a simple account of regularity and semi-regularity. For a regular verb such as *laver*, only the imperfective stem needs to be stipulated in the lexicon;

[13] If we take into account the 12 slots with a maximum of 6 stems per verb, we end up with 3,403,127 distinct ways of filling the French stem space.

[14] This proposal is parallel to the generalization that in English, the simple past is derived from the present and the past participle is derived from the simple past (Zwicky, 1991), with the difference that the relation we establish is between stems and not between inflectional forms. It is not clear from Morin's paper whether he intends the dependency relations to be relations between stems or inflectional forms.

	imperf./prst. 12pl	prst. 3pl	prst. sg
laver	lav	lav	lav
'wash'	A	A	A
savoir	sav	sav	sɛ
'know'	A	A	B
*	*	*	*
	A	B	A
mourir	mur	mœr	mœr
'die'	A	B	B
boire	byv	bwav	bwa
'drink'	A	B	C

Table 5: Dependencies between stem slots in the present paradigm

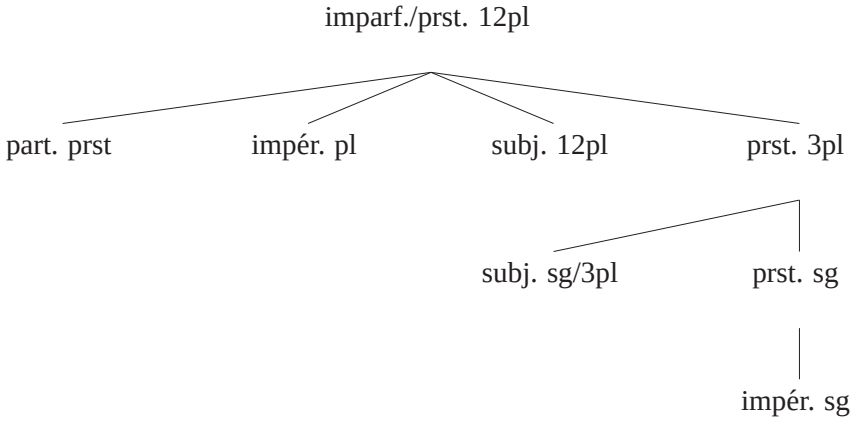


Figure 1: The stem dependency tree

the default dependencies predict that in the absence of other information, identical stems must be used for the present third plural, and for the present singular. Semi-regular verbs are simply verbs which stipulate two distinct stems in the lexicon; *savoir* stipulates the imperfective and the present singular, and *mourir* stipulates the imperfective and the present third singular. Finally, an irregular verb such as *boire* stipulates three distinct stems for the three slots.

Moreover, the dependency relations allow one to account for the gap in (5). If the dependency relations listed in (5) are used to fill the stem space, the only way to get the ABA pattern is to list the same stem twice, as an imperfective and as a present singular stem. Thus an ABA verb would not be a semi-regular verb, but a completely irregular verb with three lexically stipulated stems, two of which happen to be identical. Although such a situation cannot be excluded, it is unlikely to occur in a stable verbal paradigm.

Extending this idea of stem dependencies to the rest of the stem space, we propose a network of dependency relations, which happens to take the form of a tree. This *stem dependency tree* for French verbal inflection is shown in figure 1; the immediate dominance relation in the tree

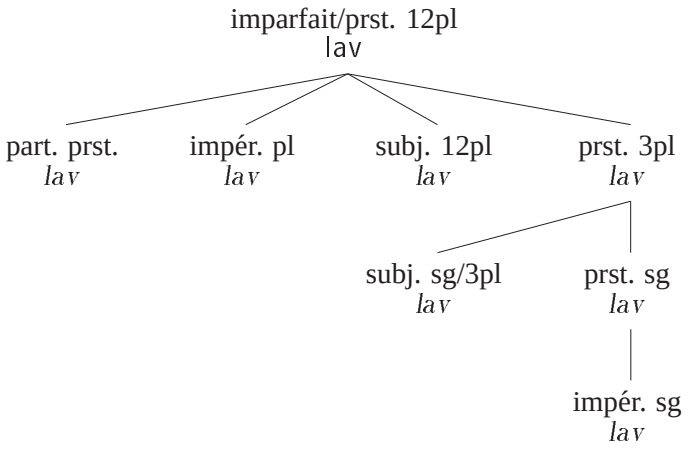


Figure 2: The stem dependency tree for *laver* (‘wash’)

represents the fact that the daughter stem is either suppletive or identical with the mother stem.

The stem dependency tree can be used by the grammar to derive predictable stems from non-predictable ones. We can assume that only those stems which are not identical to their mother in the tree need to be listed explicitly in the lexicon, and that the grammar uses the geometry of the tree to derive all the other stems. Figures 2 to 4 illustrate this with three distinct verbs: the regular *laver* (‘wash’), and two irregular verbs with different conjugation patterns, *mourir* (‘die’) and *boire* (‘drink’).[15] For the regular *laver* and all other regular verbs, only the imperfective stem needs to be listed explicitly, and all other stems are predicted to be identical. For *mourir*, two stems are listed explicitly: the imperfective, and the present third plural. This is sufficient to ensure that the imperfective, the present participle, the imperative plural and the subjunctive plural will be identical; the present third plural is different, but identical to the subjunctive, present and imperative singulars. Finally, for *boire*, three distinct stems are listed explicitly, accounting for the correct conjugation pattern.

In the preceding examples, the stem dependency tree gives us a way of avoiding redundancy in the lexical representation: each phonologically distinct stem is listed only once in the lexicon, and the tree allows the grammar to fill the stem space. Moreover, this model has the advantage of allowing only for a very restricted number of ways of filling the space: only 29 distinct ways are allowed. The fact that only 10 of the 29 distinct conjugation patterns we predict are actually attested can reasonably be taken to be a historical accident. We are thus meeting the second design requirement stated in the introduction: the analysis not only accounts for the existence of suppletion, but also for the fact that suppletion does not occur erratically in arbitrary zones of inflectional paradigms.

[15] In the figures, stems that need to be listed explicitly in the lexicon are typeset in a sans-serif font, whereas stems that are derived by the grammar are typeset in seriffed italics.

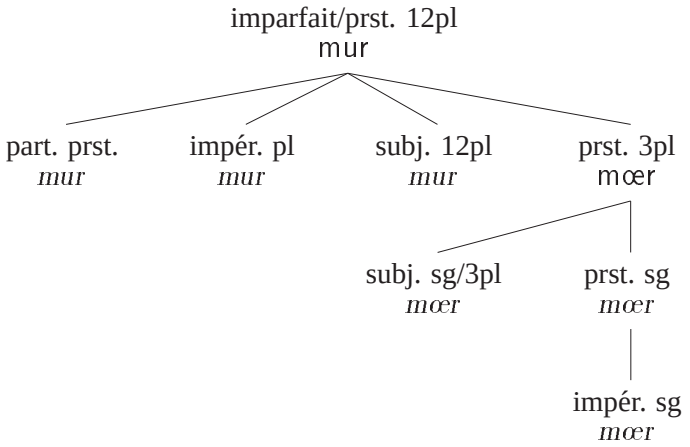


Figure 3: The stem dependency tree for *mourir* (‘die’)

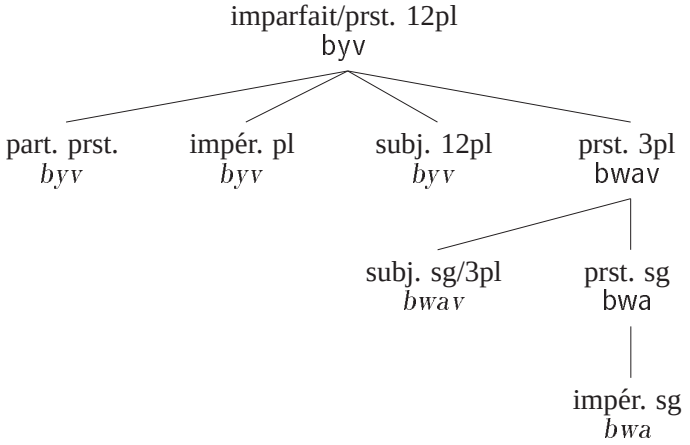


Figure 4: The stem dependency tree for *boire* (‘drink’)

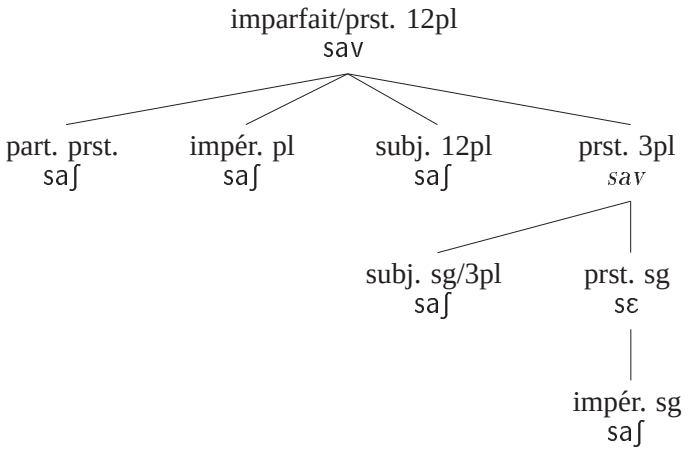


Figure 5: The stem dependency tree for *savoir* (‘know’)

1.5 Complex stem dependency

The stem dependency tree allows one to reduce drastically the quantity of redundant phonological information which has to be listed in the lexicon. However, there are still 15 verbs for which we have to list some stem more than once. An example of such a verb is the verb *savoir* (‘know’). As figure 5 shows, the geometry of the tree does not allow to predict that *savoir* uses the same stem saʃ in five different slots.

Since only 15 verbs are concerned, and nothing in the dependency model entails that no verb can list a stem twice, we could assume that this is just an accident, and that some verbs just happen to list the same stem more than once. However, once again we observe that the distribution of the redundant stems in the tree is not random, which is what we would expect if it was an accident. We observe three distinct patterns exemplified in (6).

(6)

être ‘be’	imperative sg = imperative pl = subjunctive 12pl = subjunctive sg/3pl
savoir ‘know’	prst. part. = imperative sg = imperative pl = subjunctive 12pl = subjunctive sg/3pl
faire ‘do’	subjunctive 12pl = subjunctive sg/3pl

The stem identities noted in (6) manifest the existence of a second type of dependency between pairs of stem slots, which we call *complex stem dependency*. These pairs of slots obey the pattern in (7):

(7) Complex stem dependency
There are pairs of stem slots $\langle X, Y \rangle$ such that if the stem filling X is suppletive (i.e., for French, X differs from its mother in the dependency tree) then the stem filling Y is identical with the stem filling X. In such a case, we say that Y has a complex stem dependency on X.

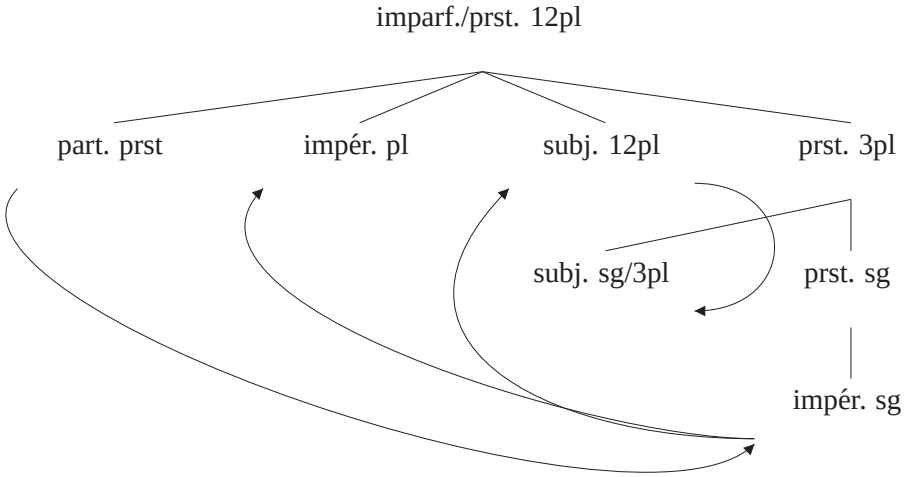


Figure 6: Enriched stem dependency tree

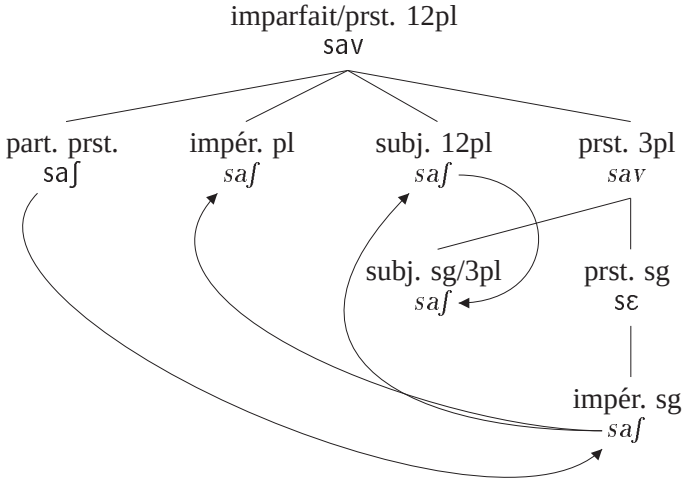


Figure 7: Enriched stem dependency tree for *savoir* ('know')

The enriched stem dependency tree presented in figure 6 exhibits the four complex stem dependencies we observe in French, which are represented as curved arrows. For instance, the imperative singular is dependent on the present participle. By taking advantage of complex dependencies, all the remaining redundant phonological information can be eliminated from the lexical description of verbs. This is shown in figure 7 for the case of the verb *savoir*. One needs only to stipulate that the present participle stem is saʃ. From this information, it can be deduced that both imperative and both subjunctive stems are saʃ. The two other patterns shown in (6) occur when instead of listing a suppletive present participle stem, a verb lists a suppletive imperative singular (in the case of *être*) or subjunctive first and second plural (in the case of *faire*).

1.6 Other types of dependency relations

Before turning to the formal analysis, we discuss how the approach to inflection discussed in this section extends to other languages and to the part of French inflectional paradigms which has been left aside.

In French, for the 8 stems we have taken into account, the dependency relation means that by default a stem is *identical* with its mother in the tree. This is not the case in general: in other languages, we observe dependency relations where the phonology of a stem is *computed* from that of its mother in the tree by default. This is what happens in English, for instance, for the dependency relation between the present and the simple past: a regular past stem is a suffixed version of the present stem. This is also what happens in a systematic way in Spanish, in Italian and in German.[16]

In fact, this situation is also found in French, in the case of the 4 stem slots which have been excluded from this discussion (future/conditional, simple past, infinitive, and past participle): these stems are not identical to the imperfective stem *even for regular verbs*; rather, regular verbs derive these stems from the imperfective stem using a non-trivial phonological function.

To account for these observations, we assume dependency relations are not (in the general case) relations of default identity, but that when a dependency relation occurs between two slots x and y, there is a single phonological function f relating the two slots. The dependency relation entails that by default, the dependent slot y can be filled with $f(x)$—the result of applying the associated phonological function to the stem it depends on.[17] It just happens that in French, the phonological function associated with most dependency relations in the stem space is the identity function on phonological strings.

Under this new, generalized view of dependency relations between morphological objects, the relation between an *inflectional form* and the stem slot it is based on can be seen as a dependency relation too: as we saw in paragraph 1.1, by default, the phonology of an inflectional form is computed from that of a stem; this default can be overridden if a verb lists a suppletive inflectional form, just as in the case of stem-to-stem relations. The only difference is that inflectional forms differ from stems in being objects that are accessible to the syntax.

To conclude this section, our analysis of inflection relies on the observation of three types of morphological dependency relations: those relating inflectional forms to their stems, and two kinds of dependency relations between stems. These relations allow us to suppress all redundant phonological information from the French lexicon, while predicting a realistic pattern of suppletion.

2 Dependency relations in the hierarchical lexicon

In this section, we provide an explicit formalization in a hierarchical lexicon of the approach to inflectional morphology outlined in the preceding section. The analysis relies on two crucial assumptions, which are discussed in the two following paragraphs. First, we represent dependency

[16] See Boyé (2000); Boyé and Cabredo Hofherr (2000) for discussion.

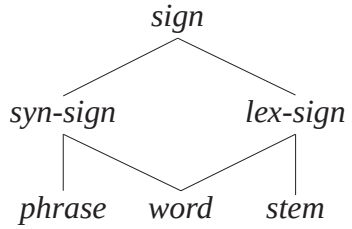
[17] This assumption works for both simple and complex dependency relation. For German and Italian, Boyé (2000) proposes complex dependency relations based on phonological functions.

relations between words and stems or among stems by assuming that lexical objects have an feature indicating other lexical signs they depend on. Second, we rely on a combination of online type construction (Kœnig and Jurafsky, 1994) and default constraints to account for suppletion.

2.1 Morphological dependencies

To integrate the analysis outlined in section 1 in an HPSG grammar, we must first assume that stems are lexical signs in their own right, distinct from words. We assume that the top of the hierarchy of signs is as in (8). *lex-sign* is the type of those signs with no syntactic structure (no DTRS feature); *syn-sign* is the type of signs which can enter syntactic composition, as stated in (9).

(8)



(9) $\textit{phrase} \Rightarrow \left[\text{DTRS} \quad \textit{list(syn-sign)}\right]$

Next, we have to account for dependency relations between lexical signs. Three types of dependency relations have been observed in section 1:

- A word can be dependent on a stem; for instance, present-3sg words are dependent on present-sg stems, in the sense that every regular present-3sg inflectional form has a phonology identical to that of the corresponding present-sg stem.
- A stem can be dependent on one other stem; for instance, present-sg stems are dependent on present-3pl stems, in the sense that every regular present-sg stem is identical to the corresponding present-3pl stem.
- A stem can be dependent on two other stems; for instance, the imperative-sg stem is dependent on both the present-sg stem and the present participle stem, in the sense that every regular imperative-sg is either identical to the present-sg (if the present participle is regular) or to the present participle (if the present participle is irregular).

To account for these relations, we assume that lexical objects take a feature M-DEP (for *morphological dependency*) which indicates for each lexical object which stems it depends on.[18] Since the morphological dependencies we are interested in relate stems from a single inflectional paradigm, and stems are underspecified for morphosyntactic features,[19] we can assume the general *Paradigm Integrity Principle* in (11):

[18]The feature M-DEP plays a role similar to that of the 'morphological daughters' feature used in some works on morphology in HPSG (Kœnig, 1999). We avoid the word 'daughters' to avoid confusion with the mother-daughter relation in the stem dependency tree, which goes in the other direction.

[19]See note 9.

(10) *lex-sign*⇒$\left[\text{M-DEP} \quad \textit{list(stem)}\right]$

(11) **Paradigm Integrity Principle**

lex-sign⇒$\left[\begin{array}{ll}\text{SYNSEM} & \boxed{1} \\ \text{M-DEP} & \textit{list}\left(\left[\text{SYNSEM} \quad \boxed{1}\right]\right)\end{array}\right]$

Morphological dependencies can then be accounted for by postulating a type in the lexical hierarchy corresponding to each of the 28 paradigm slots of a verb and to each of the 8 slots of the stem space.[20]

Types for inflectional forms specify which stem each inflectional form depends on, and specify how the phonology of the inflectional form depends on the phonology of the stem. For instance, if we disregard irregular inflection for the time being, we can assume the constraint in (12a), stating that words in the present-1pl are dependent on imperfective stems and that their phonology is that of the imperfective stem plus the suffix ɔ̃. This accounts correctly for the form of *lavons* (12b).

(12) a. *prst-1pl-wd*⇒$\left[\begin{array}{ll}\text{PHON} & \boxed{1}\oplus\tilde{\text{ɔ}} \\ \text{M-DEP} & \left\langle\left[\begin{array}{ll}\textit{imperf-prst12pl-stm} \\ \text{PHON} & \boxed{1}\end{array}\right]\right\rangle\end{array}\right]$

b. The prst. 1pl word *lavons* ‘(we) wash’

$\left[\begin{array}{ll}\text{PHON} & \text{lav}\tilde{\text{ɔ}} \\ \text{M-DEP} & \left\langle\left[\begin{array}{ll}\textit{imperf-prst12pl-stm} \\ \text{PHON} & \text{lav} \\ \text{M-DEP} & \textit{elist}\end{array}\right]\right\rangle\end{array}\right]$

For stem types, three cases must be distinguished. First, the imperfective stems are particular in being dependent on no other stems (since they are the root of the stem dependency tree). This means that each individual stem will have to be listed in the lexical hierarchy, as we will see in section 2.2.

(13) *imperf-prst12pl-stm*⇒$\left[\text{M-DEP} \quad \textit{elist}\right]$

Second, stems exhibiting a simple stem dependency are dependent on a single stem. Disregarding again cases of suppletion, these stems have the same phonology as that of the stem they depend on. This is illustrated in (14) for the present-3pl stem.

(14) *prst-3pl-stm*⇒$\left[\begin{array}{ll}\text{PHON} & \boxed{1} \\ \text{M-DEP} & \left\langle\left[\begin{array}{ll}\textit{imperf-prst12pl-stm} \\ \text{PHON} & \boxed{1}\end{array}\right]\right\rangle\end{array}\right]$

[20]Remember that for the sake of clarity, in this paper, we take into account only 8 of the 12 stems and thus 28 of the 48 inflectional forms of French verbs.

Finally, we have to account for cases of complex stem dependencies. In such cases, two stems are present on the dependency list. By convention, the first stem will always be the mother in the dependency tree. The phonology of the stem is then computed using a function choose-stm, which is defined in (15). As an example of a stem exhibiting a complex dependency, (16) gives the constraint on imperative singular stems. This constraint states that the phonology of the imperative singular is identical to the phonology of the present participle if the present participle is irregular (that is, does not have the same phonology as the stem it depends on, which is the imperfective stem); otherwise it is identical to the phonology of the present singular.

(15) choose-stm$(x, y, y) = x$
If $y \neq z$, then choose-stm$(x, y, z) = y$.

(16) *imperative-sg-stm*⇒
$$\begin{bmatrix} \text{PHON} & \text{choose-stm}(\boxed{1},\boxed{2},\boxed{3}) \\ \text{M-DEP} & \left\langle \begin{bmatrix} \textit{prst-sg-stm} \\ \text{PHON} \quad \boxed{1} \end{bmatrix}, \begin{bmatrix} \textit{prst-part-stm} \\ \text{PHON} \quad \boxed{2} \\ \text{M-DEP} \quad \begin{bmatrix} \text{PHON} \quad \boxed{3} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

2.2 Regular and irregular inflection

The constraints stated so far do not leave room to any inflectional irregularities. To take irregularities into account, we need a way of restricting the constraints stated in the preceding paragraph to regular inflection. The strategy we propose is to use a combination of online type construction and default constraints on phonological values, which is detailed in this paragraph.

2.2.1 Lexemes, stems and words

Intuitively, an irregular inflectional form is a form which has to be listed explicitly in the lexicon, because it has an unpredictable phonology; whereas a regular inflectional form is predictable from information which is already present in the lexicon. The same intuition applies to regular and irregular stems.

To capture this intuition in a hierarchical lexicon, we assume the lexical architecture shown in figure 8. All lexical signs are classified along two dimensions: LEXEME and INFLECTIONAL TYPE. We thus take lexemes to be objects which are distinct from stems: lexemes are non-maximal types in the hierarchy, with no definite phonological properties, which can be instantiated as words or stems; their role is to specify in a single type the syntactic and semantic features of a verb which are shared by all its morphological instances, even when these are not related phonologically. Stems are maximal types in the hierarchy, which have a definite phonology, and which are just as concrete as words—where they differ with words is in not being objects accessible to syntax; their role in the analysis is to ground the 'derivation' of the phonological form of individual words.[21]

[21]Thus lexemes and stems play in our analysis two roles which are often conflated in HPSG analyses—e.g. in

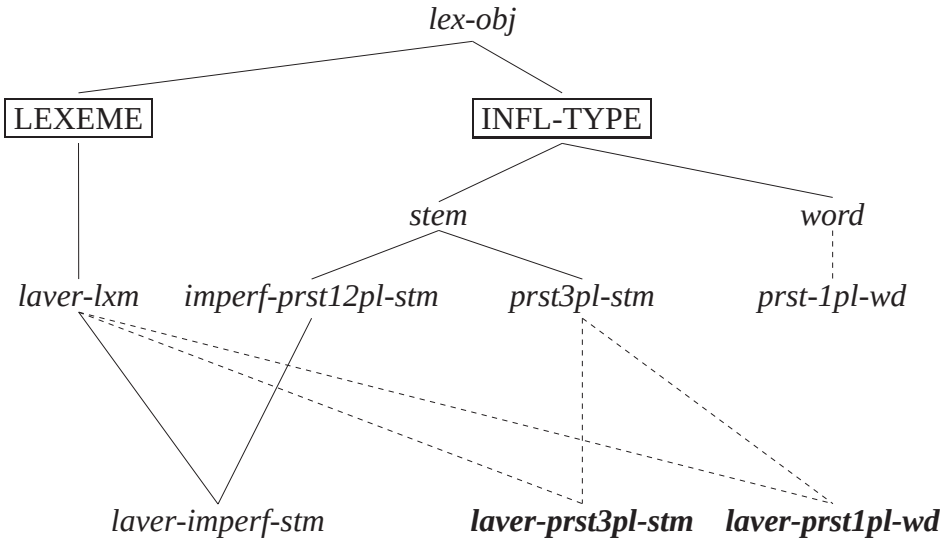


Figure 8: Lexemes, stems and words

The advantage of this two-way distinction between lexemes, stems and words is that we can now assume that a lexical sign is listed explicitly in the lexical hierarchy if and only if it is irregular; if it is regular, it will be derived by online type construction. Since the imperfective stems are at the root of the stem dependency tree, they are all irregular in the sense that they have an unpredictable phonology. Thus they are all listed explicitly as subtypes of their lexemes. For other inflectional types, regular lexical signs need not be listed explicitly, and can be deduced from the structure of the hierarchy.[22]

2.2.2 Blocking

The contrast between regular and irregular inflection is illustrated in more detail in figure 9 for present-1pl inflectional forms. Since the verb *être* ('to be') has an irregular present-1pl, this word is listed explicitly as a subtype of both *être-lxm* and the inflectional class *prst-1pl-wd*. Since *laver* ('to wash') is regular, no present-1pl inflectional form is listed explicitly for this verb; but we know from the structure of the hierarchy that there exists a maximal type that is both a subtype of *laver-lxm* and of the inflectional class *prst-1pl-wd*. Online type construction gives us a simple account of (inflectional) morphological blocking: there is no 'regularized' present-1pl form for *être*, simply because the slot this form would fill in the hierarchy is already filled by the irregular *sommes*.[23]

(Kœnig, 1999) or (Sag and Wasow, 1999). One advantage of the lexical architecture of figure 8 is that it integrates straightforwardly with the analysis of the *semantics* of inflection proposed in Bonami (2001), thus allowing a unified view of inflection.

[22]In figure 8 and later figures, types constructed by online type construction are highlighted by being typeset in boldface and related to their supertypes by dashed lines.

[23]Notice that our approach to blocking differs from most other approaches in feature-structure-based theories in not relying on a comparison of morphosyntactic features of concurrent lexical entries (Andrews, 1990; Blevins, 1995; Briscoe et al., 1995). Rather, it is purely based on the notion of regularity.

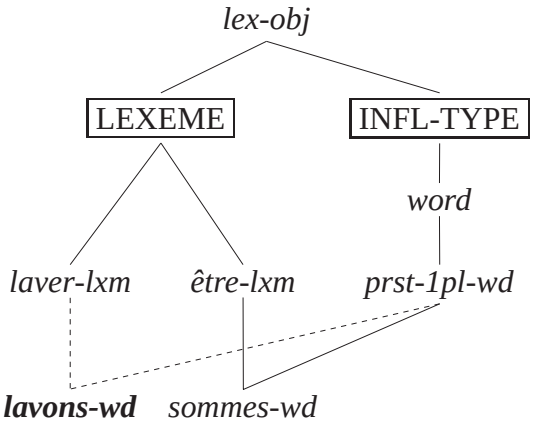


Figure 9: Present-1pl inflectional forms

2.2.3 Default phonologies

It is a crucial feature of our analysis that irregular forms are subtypes of the inflectional types they instantiate—for instance, *sommes* is a subtype of the present-1pl word type. A consequence of this hypothesis is that the constraints on the phonology of dependent inflectional types proposed in paragraph 2.1 cannot be maintained; for instance, it is not the case in general that present-1pl words end in /5̃/ (12). For the system to work, we need to re-state the constraints on the phonology of inflectional forms as default constraints. For instance the constraint in (12) must be stated as (17), where the phonological dependency of the word on the imperfective stem is only true by default.

(17) *prst-1pl-wd*⇒
$$\begin{bmatrix} \text{PHON} & /\boxed{1}\oplus 5̃/ \\ \text{M-DEP} & \left\langle \begin{bmatrix} \textit{imperf-prst12pl-stm} \\ \text{PHON} \quad \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$$

The default specification in (17) is a direct reflection of the original notion of morphological dependency discussed in section 1: morphological dependencies constrain the phonology of regular forms, but do not say anything about the phonology of irregular forms.

2.2.4 The case of dependent stems

The approach to regularity and irregularity just outlined in the case of inflectional forms applies directly to the case of dependent stems. We now illustrate the analysis in figure 10 with the case of the *prst-3pl-stm*. As we saw before, this stem type is simply dependent on the imperfective stem type.[24] The constraint on the present-3pl stem thus states that by default, its phonology is the same as that of the imperfective stem. This is enough to make sure that a regular stem such as *laver-prst3pl-stm* will inherit the phonology of the corresponding imperfective stem. But of course, an

[24]Stems exhibiting a complex dependency have the same analysis, except that the (default) phonological specification makes use of the choose-stm function defined in (15).

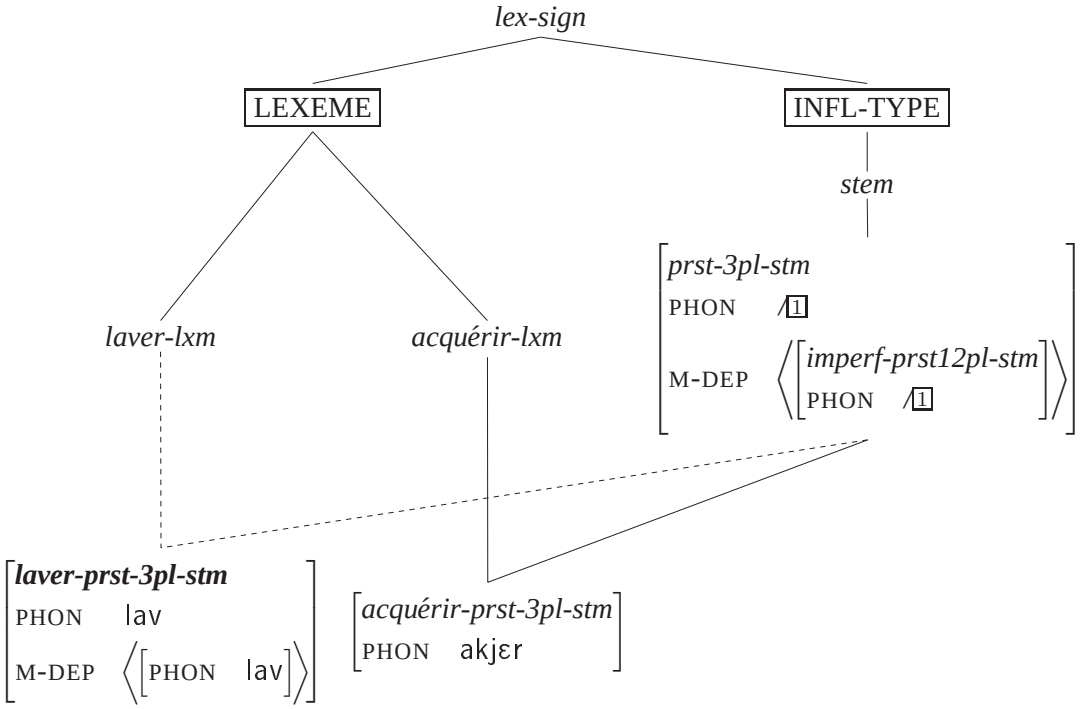


Figure 10: Present-3pl stems

irregular stem will override the default phonology. This is illustrated here with the stem of the verb *acquérir* (‘to acquire’).

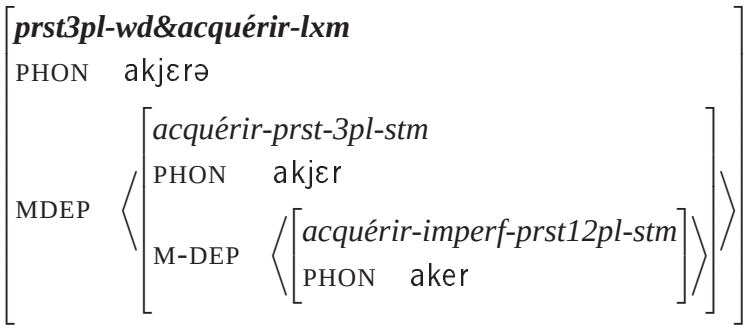
The two stems in figure 10 can then be used to form regular present-3pl words, as shown in (18). (18a) is the constraint on the *prst-3pl-wd* type, and states that words of this type suffix by default a schwa to the stem they depend on. Both words in (18b–c) are regular, but (18b) is dependent on the regular *laver-prst-3pl-stm* whereas (18c) is dependent on the irregular *acquérir-prst-3pl-stm*.

(18) a. *prst-3pl-wd*⇒
$$\begin{bmatrix} \text{PHON} & /\boxed{1}\oplus\text{ə} \\ \text{M-DEP} & \left\langle \begin{bmatrix} \textit{prst-3pl-stm} \\ \text{PHON} \quad \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$$

b. The prst. 3pl word *lavent* ‘(they) wash’

$$\begin{bmatrix} \textbf{\textit{prst3pl-wd\&laver-lxm}} \\ \text{PHON} \quad \text{lavə} \\ \text{MDEP} \quad \left\langle \begin{bmatrix} \textbf{\textit{laver-prst-3pl-stm}} \\ \text{PHON} \quad \boxed{1}\text{lav} \\ \text{M-DEP} \quad \left\langle \begin{bmatrix} \textit{laver-imperf-prst12pl-stm} \\ \text{PHON} \quad \boxed{1}\text{lav} \end{bmatrix} \right\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

c. The prst. 3pl word *acquièrent* ‘(they) acquire’



3 Summary

In this paper, we have given evidence that inflectional irregularities are not erratic: there is an organization of the irregular forms, which is adequately represented by postulating a stem space structured by morphological dependency relations. Although this paper has focussed on French, it is clear from the case studies in (Boyé, 2000) that different languages will use different dependency relations, resulting in dependency trees with language-particular structures. The stem slots that are needed to model irregular inflection appear to be purely morphological objects, deprived of morphosyntactic features—*morphomes* in the sense of Aronoff (1994): there is no correlation between the classification of inflectional forms operated by the stem space and the morphosyntactic classification operated by tense and agreement features.

The second part of the paper has shown that the view of inflection presented in the first part can be made more precise in an HPSG hierarchical lexicon. We have presented a novel approach to (inflectional) morphological blocking which relies on both online type construction and default constraints on phonology. The advantage of this approach is that it expresses directly the notion of inflectional suppletion which was shown to be appropriate in the first section: a sign is suppletive if and only if its phonology is unpredictable.

References

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

Andrews, A. D. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory* 8:507–557.

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

Blevins, J. P. 1995. Syncretism and paradigmatic opposition. *Linguistics and Philosophy* 18:113–152.

Bonami, O. 2001. A syntax-semantics interface for tense and aspect in French. this volume.

Boyé, G. 2000. *Problèmes de morpho-phonologie verbale en français, espagnol et italien*. Ph.D. thesis, Université Paris 7.

Boyé, G. and P. Cabredo Hofherr. 2000. Spanish verb conjugation: towards an analysis without groups. ms.

Briscoe, T., A. Copestake, and A. Lascarides. 1995. Blocking. In P. Saint-Dizier and E. Viegas, eds., *Computational lexical semantics*. Cambridge: Cambridge University Press.

Fradin, B. 1993. *Organisation de l'information lexicale et interface morphologie/syntaxe dans le domaine verbal*. Ph.D. thesis, Université Paris 8.

Kœnig, J.-P. 1999. *Lexical relations*. Stanford: CSLI Publications.

Kœnig, J.-P. and D. Jurafsky. 1994. Type underspecification and on-line type construction. In *Proceedings of WCCFL XIII*, pages 270–285.

Lieber, R. 1980. *On the organization of the lexicon*. Ph.D. thesis, MIT, Cambridge, MA.

Matthews, P. H. 1972. *Inflectional morphology*. Cambridge: Cambridge University Press.

Morin, Y.-C. 1987. Remarques sur l'organisation de la flexion en français. *ITL Review of Applied Linguistics* 77–78:13–91.

Plénat, M. 1984. Recherches sur les catégories verbales (et autres). Études de morphologie française. Thèse de Doctorat d'état, Université de Vincennes à Saint-Denis.

Plénat, M. 1987. Morphologie du passé simple et du passé composé des verbes de l' "autre" conjugaison. *ITL Review of Applied Linguistics* 77–78:93–150.

Sag, I. A. and T. Wasow. 1999. *Syntactic theory. A formal introduction*. Stanford: CSLI Publications.

Selkirk, E. 1982. *The syntax of words*. Cambridge: MIT Press.

Spencer, A. 1991. *Morphological Theory: an introduction to word structure in generative grammar*. Oxford: Basil Blackwell.

Stump, G. T. 1993. On rules of referral. *Language* 69(3):449–479.

Swiggers, S. and K. Van Den Eynde. 1987. La morphologie du verbe français. *ITL Review of Applied Linguistics* 77–78:151–251.

Van Den Eynde, K. and C. Blanche-Benveniste. 1970. Essai d'analyse de la morphologie du verbe français. Présentation d'hypothèses de travail. *Orbis* XIX(2).

Zwicky, A. M. 1985. How to describe inflection. *Berkeley Linguistics Society* 11.

Zwicky, A. M. 1991. Systematic versus accidental phonological identity. In G. Bossong and B. Comrie, eds., *Paradigms: The Economy of Inflection*, pages 113–131. Mouton de Gruyter.