RECHERCHES LINGUISTIQUES DE VINCENNES

Recherches linguistiques de Vincennes

39 | 2010
Racine et radical

The lexical representation of nouns and ajectives in romance languages

Fabio Montermini

OpenEdition Journals

Electronic version
URL: http://journals.openedition.org/rlv/1869
DOI: 10.4000/rlv.1869
ISSN: 1958-9239

Publisher
Presses universitaires de Vincennes

Printed version
Date of publication: 1 December 2010
Number of pages: 135-162
ISBN: 978-2-84292-264-1
ISSN: 0986-6124

Electronic reference
Fabio Montermini, « The lexical representation of nouns and ajectives in romance languages », *Recherches linguistiques de Vincennes* [Online], 39 | 2010, Online since 01 December 2014, connection on 21 April 2019. URL : http://journals.openedition.org/rlv/1869 ; DOI : 10.4000/rlv.1869

Recherches Linguistiques de Vincennes 39 – 2010 — p. 135-162

Fabio Montermini
CNRS – Université de Toulouse-Le Mirail

THE LEXICAL REPRESENTATION OF NOUNS AND ADJECTIVES IN ROMANCE LANGUAGES

Abstract

This paper discusses such issues as the format under which lexical units are stored in memory, the way in which inflection and derivation interact, and the definition of morphological units, such as stem, root, word form, etc. More largely, two competing models of morphological competence, a "units-plus-rules" model and a global model based on analogy, are discussed. The main focus is on nouns and adjectives in three Romance languages: French, Catalan and Italian. The data are analyzed within a word-based model of morphology. Word forms are considered to be the main units of lexical memorization universally. These forms are organized into larger, more abstract units, the lexemes, roughly corresponding to dictionary entries. It is claimed that the other units traditionally identified in morphological analyses, such as stems, roots, etc. may serve as organizing units in some cases, but should not be considered to be universal units of morphology cross-linguistically.

Keywords

Morphology, lexicon, allomorphy, paradigms, Italian, French, Catalan.

The main goal of this paper[1] is to discuss such issues as the format under which lexical units are stored in memory, the way in which inflection and morphological derivation interact, and the very definition of morphological units, such as "stem", "root", "word form", etc. More largely, two competing models of morphological competence, a 'units-plus-rules' model and a global model based on analogy, will be discussed.

The main focus is on nouns and adjectives (the label 'nominals' encompasses members of both categories) in three Romance languages: French, Catalan and Italian. These three languages display peculiar characteristics in relation to nominal inflection. In particular, French nouns and adjectives are almost always monomorphemic under a morphemic analysis; some Catalan nouns and adjectives are unmarked, while others display overt inflectional markers; finally, Italian nominals systematically display inflectional markers, *i.e*. they are never monomorphemic under a morphemic analysis.

I claim that the data observed are better explained within a word-and-paradigms realizational model of morphology (*cf.* Stump, 2001). However, most of the analyses I propose are compatible, via some readjustments, with different, rule-based, models. In particular, the model of morphology I defend is a word-based one, as proposed by Blevins (2006). That means that I consider word forms as the main units of lexical memorization universally. These forms are organized into larger, more abstract units, the lexemes, roughly corresponding to dictionary entries. I consider that the other units traditionally identified in morphological analyses, such as stems, roots, etc. may serve as organizing units in some cases, but should not be considered to be universal units of morphology cross-linguistically. I also adopt a so-called "thematic" model of morphology (*cf.* Bonami & Boyé, 2007 for a presentation; Plénat, 2008, 2009 for some applications to French). In this model, the identification of a single abstract form underlying the various inflected forms of a lexeme is a secondary issue. Rather, the model in question focuses on the network of relationships between forms and on the corresponding patterns, which are recurrent among lexemes. Forms that are linked by a phonological rule and forms that are linked by an apparently arbitrary relationship are not different in nature. This approach avoids the multiplication of lexical-specific or category-specific rules. Such rules are often proposed to account for relationships among forms which are simply the reflect of phonological processes that were active in the past, but are now completely lexicalized or morphologized.

This paper is organized as follows: in section 1 I discuss some models of morphology, focusing in particular on the kind of units that are considered as the

1. I thank Florence Villoing for her incitation to work on stems. I am also grateful to Jesse Tseng and to the anonymous reviewers of this paper, whose valuable comments helped me to improve this paper. I am, of course, responsible for all the remaining errors.

basic objects stored in a speaker's lexicon. In section 2 I present the advantages of a word-based thematic approach, in particular for fusional languages, and propose an analysis of French data. Section 3 presents and analyzes the Catalan and Italian data. Finally, section 4 contains some concluding remarks.

1. Models of morphological competence: an overview

1.1. The storage of lexical units

The relationship between units that are built on-line and units that are memorized is a recurrent topic in morphological and more globally in linguistic studies. Models of the representation of complex morphological units (inflected or derived words) may be situated between two poles, and are in general divided into three categories:

- so-called "impoverished entry" theories (*cf.* Jackendoff, 1975: 642), according to which the lexicon only contains idiosyncratic information, *i.e.* those features of complex words that cannot be deduced by rule. According to these theories, the lexicon only contains irregular forms (*e.g.* the plurals *oxen* or *feet*), not regular ones (*e.g.* *books*), and is fundamentally seen as an unstructured list of exceptions. This conception of the lexicon, which goes back to the structuralist tradition, has been mainly supported by formal linguistics, in particular within the generative framework (*cf.* Halle, 1973 for its introduction into morphology). It has also been adopted in a series of studies in psycholinguistics and cognitive linguistics, for instance in the well known "Dual Route Model" (Pinker & Ullman, 2002);
- "full entry" theories, according to which all words are stored in the lexicon, independent of their constitution (*cf.* Bybee, 1985; Blevins, 2006).
- mixed models, according to which the lexicon stores at least all idiosyncratic information, but that it is not necessary to think that "the lexicon stores information non-redundantly" (Jackendoff, 2002: 153).

In both full entry and mixed models, the lexicon is seen as a structured entity and is organized by a network of relations between the forms, for instance by redundancy rules (*cf.* Jackendoff, 1975).

One of the most serious problems posed by impoverished theories is that they correspond to a theoretical desideratum, but they appear to be the most implausible ones, from a cognitive and a psychological point of view. In fact, there seems to be no psycholinguistic evidence of the fact that the human brain needs to store only non-redundant information, whereas there is rich evidence of the fact that some regular complex words (at least the most frequent ones) are stored in the lexicon along with irregulars (*cf.* Stemberger

& McWhinney, 1986; Bybee, 2007: 207-208; Baayen *et al.*, 2003, among others). Moreover, as several studies have pointed out (see *e.g.* Derwing & Skousen, 1988), impoverished models are only apparently more economic in cognitive terms, since a poorer lexicon implies a richer and more complex system of rules. However, nothing indicates that in linguistic ability priority must be given to computation rather than to retrieval of stored information, and psycholinguistic evidence, once again, seems to prove the contrary (*cf.* Derwing & Skousen, 1988: 60; Baayen, 2003).

On the other hand, the distinction between full-entry or mixed models presented above is probably a false problem. In any case, it cannot be satisfactorily resolved until we have given an answer to two questions, which are rarely explicitly addressed in the morphological literature, namely what is the lexicon, and what is a regular form.

Concerning the first question, there are at least two definitions of the lexicon, although, unfortunately, scholars do not always make the distinction[2]. "Lexicon" may refer to the actual mental lexicon of each speaker or to an idealized entity. Access to the first is of course impossible and probably not so useful; what we can hope to have an idea of is the average lexicon of an average speaker. However, we can only achieve an imperfect representation of this idealized notion of the lexicon, and in particular we can hope to have a clear picture of its centre, but we should probably abandon the goal of precisely defining its borders.

As far as regularity is concerned, this notion is probably better described as a gradient than as a binary distinction. If this is true, it is impossible to draw a clear dichotomy between two objects that are distinct in nature, with on the one hand regular forms built and processed by computation and on the other hand irregular forms retrieved from a list of idiosyncrasies. The existence of subregularities and attraction phenomena within limited sets of words is a fact that is commonly observed in many languages. In fact, the identification of regular and irregular classes of items in a language is probably more a matter of language design, and possibly of metalinguistic tradition, than a matter of substance. Let us take a simple example. In the majority of Romance languages, nouns and adjectives may be divided into several inflectional classes. These classes may be considered as "regular", since all the lexemes they contain display an identical behaviour in inflection. For instance, in Italian, there are three such classes of adjectives: the two largest ones contain adjectives with four forms and two forms; there is also a smaller class of adjectives in *-a*/*-i* (mostly including words containing a suffix of Greek origin derived from nouns), and some invariable adjectives. (1) summarizes the situation, and

2. For an overview of the different meanings of the word "lexicon", see Aronoff (1994: 16-22).

Table 1 below gives a rough idea of the distribution of the classes in the Italian lexicon (figures are based on a count from a corpus of spoken Italian, De Mauro *et al*., 1993):

(1)

	singular		plural	
	masc.	fem.	masc.	fem.
-o class (‘beautiful’)	bello	bella	belli	belle
-e class (‘strong’)	forte		forti	
-a class (‘idiot’)	idiota		idioti	idiote
invariables (‘pink’)	rosa			

-o class	1,853
-e class	855
-a class	44
invariables	17

Table 1: distribution of adjective classes in Italian

Italian is an example of a language in which adjectives may be divided into at least two homogeneous classes, with a relatively small number of idiosyncratic cases. On the other hand, in French the default case for an adjective is to be (phonologically) invariable (*cf. pur*, *purs*, *pure*, *pures*, all [pyr], ‘pure’), but there are some adjectives which vary, mostly for gender (*cf.* masculine *beau, beaux* [bo] *vs* feminine *belle*, *belles* [bɛl], ‘beautiful’). There is some debate if the alternations found in French adjectives should be treated as rules (for instance based on phonological principles) or as idiosyncratic (*cf.* Tranel, 1981; Bonami & Boyé, 2005, for an overview). In favour of the second hypothesis there is the fact that the relation between the two forms is not always predictable (*cf. lent* [lã] / *lente* [lãt] *vs grand* [grã] / *grande* [grãd] ‘big’) and that it is not phonologically conditioned. Table 2 presents the distribution of the default and of some of the most recurrent patterns for French adjectives (based on the data in Lexique.3, http://www.lexique.org).[3]

3. An anonymous reviewer proposed such neologisms as *netteux* [nɛtø] / *netteuse* [nɛtœz] (‘netaholic’) or *amusard* [amyzar] / *amusarde* [amyzard] (‘party animal’) as possible counterexamples to the figures given in Table 2. Of course, the statistics I present do not distinguish between simple and derived adjectives. Derived adjectives are inflected according to the variation pattern instantiated by their suffix, which, in this case too, may be invariability (*cf. désirable* / *désirable* ‘desirable’) or a more or less recurrent pattern of variation (*cf. désireux* [dezirø] / *désireuse* [dezirœz] ‘desirous’).

invariables	*pur – pure* 'pure'	80.2%
ã ~ ãt	*lent – lente* 'slow'	7.1%
ø ~ œz	*heureux – heureuse* 'happy'	2.8%
ɛ̃ ~ ɛn	*sain – saine* 'healthy'	2.3%
if ~ iv	*juif – juive* 'Jew'	1.6%
e ~ ɛr	*léger – légère* 'light'	0.9%
ar ~ ard	*bavard – bavarde* 'talkative'	0.7%
ɛ̃ ~ in	*fin – fine* 'fine'	0.6%
others		3.8%

Table 2: distribution of adjective classes in French

The situation of French adjectives is much more similar to that of English verbs, with a large default class and a smaller number of units displaying heterogeneous behaviours. However, some of the subgroups listed in Table 2 seem to correspond to the default for particular subclasses. For instance, when the masculine of a French adjective ends with a nasal vowel, invariability is very rare and limited to special classes of items (for instance, colour terms), while other patterns, such as denasalization or the adjunction of a consonant (mostly [t] or [d]), are much more frequent. The relevant figures (data from Lexique.3) are given in Table 3:[4]

	masculine in		
	ã	ɛ̃	ɔ̃
invariable	4	4	7
denasalization	36	493	46
+t	1141	26	1
+d	14	–	12
+C (other than t or d)	1	–	3

Table 3: inflection patterns for adjectives ending in a nasal vowel in French

As it can be seen, the notion of regularity is relative for French adjectives. When a speaker of French has to create the feminine form of an unknown / novel adjective on the basis of the masculine, the default choice is identity. Of course, this is true when no other parameter intervenes, for instance when the unknown adjective does not contain a derivational suffix, or a suffix-

4. I consider a variety of French in which the distinction between [ɛ̃] and [œ̃] is neutralized. For masculine adjectives in [ɛ̃] denasalization includes feminine forms in [ɛn], [in] and [yn].

like sequence, and when it does not contain a sequence which is frequent in alternating adjectives. For instance, adjectives constructed in *verlan* are very often invariable (*e.g. relou* < *lourd* 'heavy', *ouf* < *fou* 'crazy').[5] On the contrary, when the adjective in question ends in a nasal vowel in the masculine, identity is virtually unavailable. More patterns (some of which are represented by a relatively small number of adjectives) have to be taken into account. I maintain nevertheless that the distinction between the patterns that are active in Italian and those that are active in French is a quantitative one, but not a qualitative one, since, as we have seen, both can serve as models for the inflection of novel / unknown lexemes. Several recent studies have shown that default inflectional patterns are sensitive to frequency and neighbourhood effects (*cf.* Baayen *et al.*, 2003; Albright, 2002), thus suggesting that analogy is a more widespread phenomenon than is commonly believed, and underlining the weight of probabilistic principles in the organization of morphology (*cf.* Baayen *et al.*, 2003; Albright, 2009; Plag, 2009, among others).

To go back to the issue of lexical storage, it is probable, under this view, that the main task of a theory of morphology is to define what "must" be stored in the lexicon, and what "can" be, but is not necessarily stored in individual lexicons; more globally, it must describe the network of relationships existing between lexical forms, and how new forms are built on the basis of this network.

1.2. Models of morphological analysis

A model of the lexicon that includes redundancy and the storage of regular units along with irregulars is more compatible with a non-symbolic and non-additive approach to morphology. In this section I sketch the main assumptions of this approach. In particular, I will propose an analysis of Italian nominal inflection within a so-called "thematic" approach to morphology, which will be presented in section 2.

The model of morphology I propose is an inferential-realizational one, according to Stump's (2001: ch. 1) classification. It is inferential in that it considers that the traditional morphemes are not objects but rather the exponents of morphological relations between forms. Accordingly, there is no difference in substance between concatenative and non-concatenative morphological processes. The adjunction of phonological material at the right or at the left edges of a base (affixation) is just the most common of morphological operations, but it is not different in nature from the other types

5. *Verlan* is a morphophonological process, frequent in non-standard youth slang, which consists, roughly, in inverting the syllables of a word (*cf.* Fradin *et al.* 2009: 28-32 for a recent overview). I thank an anonymous reviewer for this suggestion.

of operations which can be observed across languages (including segmental and suprasegmental operations, such as apophony, reduplication, stress or tone shifts, etc.). This is consistent with the observation that, in inflection and in word formation, the same morphosyntactic and semantic values may be expressed either by affixation or by any other morphological means with no real difference. It is realizational in that, in line with what I have just described, it considers that, for instance, the morphosyntactic features of an inflected form are expressed ("realized") by the form as a whole, and not by some of its subparts (for instance by an affix). Under this view, the inflectional make up of a lexeme (its paradigm) is prior to the actual forms that realize it and partially unconnected with them. In other terms, in a form like *books* the value [plural] for the feature Number is not carried by the morpheme *-s*, as in structuralist and much of generative morphology, but rather, it is the fact that the lexeme BOOK, as a noun, is intended to have a plural that determines the fact that a [s] is adjoined to its base, but this particular value may be realized, for other lexemes, by another form (*e.g. oxen* or *feet*).

Additive (or incremental, according to Stump's terminology) models of morphology consider, on the contrary, that morphemes are linguistic objects, possibly stored in the lexicon along with autonomous words, and that the construction of meaning is realized step by step: ideally, each portion of phonological material that is adjoined to a base adds a portion of meaning. This model is particularly suitable for dealing with a limited number of phenomena in a limited number of languages, namely the agglutinating ones. However, the problems encountered by morphemic approaches for other types of phenomena and of languages (extended or cumulative exponence, zero marking, subtraction, etc.) are well known and documented.[6] Of course, there have been several attempts to save morphemes in these cases, for instance by recurring to "special" morphemes, but most of these explanations have an *ad hoc* character and such units are posited merely for theoretically internal reasons. From an empirical point of view, morphemic and non-morphemic approaches are equally suitable for dealing with purely agglutinative phenomena; the latter, however are empirically more adequate for dealing with non-agglutinative phenomena, and should therefore be preferred.

The dereification of morphemes has several desirable effects. First, as already observed, it allows treating concatenative and non-concatenative phenomena in the same manner, namely as exponents of morphological operations, and it makes a truly output-oriented approach to morphology possible. Second, actually observed morphologically complex words are not evaluated in binary terms, as "grammatical" or "ungrammatical", but rather

6. For detailed criticisms of the morphemic approach to morphology *cf.* in particular Matthews (1972: ch. 6); Stump (2001: ch. 1); Fradin (2003: ch. II).

with relation to the optimal output that can be expected from the interaction between a base and the operation performed on it. This approach has obvious connections both with Optimality-Theoretic[7] models and with the canonical approach to inflection, developed in particular by Corbett (2007), and is clearly compatible with the probabilistic view of morphology referred to in the previous section. Finally, in a morphology where the morphemes do not have an existence *per se*, the exact decomposition of complex words and the identification of the frontier between a base and an affix are no longer an issue, and affixal allomorphy can be viewed as a much less complex phenomenon. To take a brief example, let us consider deverbal action nouns in Italian, an issue that has been addressed by several authors.[8] Some data are given in (2) (*cf.* Montermini, 2006 for more data and for details):

(2)

verb	past participle	derived form	
interrogare	interrogato	interrogazione	'interrogation'
estinguere	estinto	estinzione	'extinction'
decidere	deciso	decisione	'decision'
riscuotere	riscosso	riscossione	'tax collection'
espellere	espulso	espulsione	'expulsion'

The only sequence which is common to all the derived forms is [jone], but there are restrictions also on the two phonemes preceding this sequence, respectively, the affricate [ts] (graphically *z*) or the sibilant [sː/z], and a vowel or a sonorant, which gives 20 possible forms for the suffix overall. As can be seen from the data in (2), the form of the derived form is systematically linked with the form of the pas participle. Clearly, the final derived form comes from the interaction between the form of the past participle (most of the verbs exemplified have unpredictable participles) and the form of the suffix. In Montermini (2006) I proposed that the form of the suffix itself may be represented as a set of hierarchically ordered constraints, as shown in (3):

(3) | a > i > e > o > u > sonorant | ts > sː/z | jone

where the form occupying the highest place in the hierarchy ([atsːjone]) corresponds to the form of the deverbal noun for the default class of Italian verbs. Under a purely morphemic analysis, we would be forced either to list 20 different allomorphs for the same morpheme, or to give a maximally underspecified representation for it (*e.g.* [jone]), thus missing the generalization

7. See in particular Xu (2007) for an implementation of realizational morphology within OT.

8. See Scalise (1983); Thornton (1990/1991); Burzio (2003), among others.

that the sequences preceding it are limited in number and in the vast majority of cases parallel to the past participle.[9]

Another property of the model of morphology I propose is that it is emergent (or "abstractive" in Blevins, 2006 terms). In such a model the rules of morphology emerge as generalizations from the existing lexicon, and the lexicon itself emerges as a generalization from the concrete linguistic expressions a speaker is exposed to. This conception of morphology is consistent with a series of studies in which the lexicon is conceived as a network of relations from which morphological patterns emerge as generalizations on the basis of the formal and semantic connections existing between words (*cf.* Bybee, 1985; Burzio, 2002; Blevins, 2006, among others). This approach tries to propose a formalization of the classic concept of analogy (*cf.* Derwing & Skousen, 1988 and, for a recent survey, Blevins & Blevins, 2009). It rejects the idea that there is a clear-cut distinction between a computational grammar and a lexicon of memorized items. Rather, the grammar is structured by the regularities and subregularities that emerge from the connections between the items memorized in the lexicon. Under this view, any relation existing between two forms in the lexicon is potentially a "rule" of morphology, and, as observed above, the distinction between different rules is rather quantitative than qualitative. What is generally called a "rule", in fact, is nothing more than a very general analogy (*cf.* Blevins & Blevins, 2009: 10).

Not only rules, but also the basic units of morphology emerge from actual linguistic productions. The next section discusses the nature and the theoretical plausibility of traditional morphological units.

1.3. The units of morphology

Morphological models are generally divided into morpheme- and lexeme-based. Clearly, the realizational and emergent model I propose takes the words, rather than the morphemes, as the basic units of morphological analysis. Theories that seek to deny the theoretical status of the word underline the difficulty of defining this notion unambiguously. However, the concept of word seems to have a psychological and cognitive salience cross-linguistically, while the same is more difficult to prove for morphemes (*cf.* Dixon & Aikhenvald, 2002 for the most detailed survey of the notion of word I am aware of). In the model of morphology I defend words are the basic unit of morphological organization that emerge from the lexicon. In particular, I think of the word as a lexeme, *i.e.* as an abstract unit that can appear in

9. An exception is constituted by some verbs in *-uire*, like *attribuire – attribuzione* ('attribute' / 'attribution'). (Thanks to Anna M. Thornton for having drawn my attention on these examples.)

syntactic constructions under different forms. A word is thus seen as a network of connected forms, and the relationships between these forms are structured into a paradigm.

Typical incremental approaches are based, as far as the lexical representation of units is concerned, on two assumptions which, although they are disputable, are never discussed, namely that: (i) the ideal lexical representation has the format one unit = one form, and (ii) this form must be the same for all units. On the contrary, in the approach I defend, the identification of a unique underlying form for a lexeme is not an issue, not even for the regular ones. This does not mean, of course, that all forms are necessarily stored in the lexicon for all lexemes. Fully regular lexemes are precisely those for which one form is generally sufficient to reconstruct the whole paradigm. As observed above, however, morphology may define what must minimally be stored in the lexicon, not what is actually stored. For some lexemes, I will claim, what must be stored does not necessarily coincide with one specific form. The identification of a form of the lexeme as its basic form is often an arbitrary and purely theory-internal choice. I will come back to this issue in section 3.2., where I will discuss the Italian data.

Words (lexemes) are thus considered as the basic units of morphological organization cross-linguistically. It is clear that the structure of their paradigm and the relations between their forms are more important, in this model, than the segmentation of individual forms and the identification of subparts. Several of these subparts, such as roots, stems, affixes, etc., are considered as theoretically relevant units in several approaches. On the contrary, I claim that such units may have a descriptive and empirical value, but that they do not correspond to units that have a theoretical status in describing the speakers' morphological competence. Affixes, as we have seen, are simply the exponents of morphological operations, *i.e.* the formal expression of a relationship between two word forms. As far as roots and stems are concerned, they are better described as the minimal common representation of a set of forms than as real morphological units with a clearly defined phonological representation. To take the examples of Italian adjectives in (1), we can easily say that the stem of ROSA is ['rɔːza], and that all the forms of this lexeme are realized by identity with its stem. With a more complex lexeme, *e.g.* BELLO, we can describe the set of functions linking all the inflected forms (as in (4)), and say that it is sufficient for a speaker to memorize one of them in order to inflect the whole lexeme. However, choosing one of the forms as "the" basic stem of the lexeme would be an arbitrary choice: each individual speaker may have memorized one or more of them, since they can all function equally well as basic:

(4) Masc Sg XV → Xo
Fem Sg XV → Xa
Masc Pl XV → Xi
Fem Pl XV → Xe

In what follows, then, I will consider "stem" as a label for a set of connected forms that systematically covary within a paradigm. For practical reasons, we may interpret the stem as the phonological sequence that allows us to obtain all the forms of the set with the smallest number of rules, but nothing indicates that it is precisely this stem that is stored in a speaker's lexicon.

To resume, I consider lexemes (abstract words) and (concrete) word forms to be the actual basic units of morphology, and I recognize the existence of an intermediate level, in which stems, roughly defined as a subparadigm, *i.e.* a set of forms systematically covarying, are the unit of organization. The principles of this lexeme-based and thematic model of morphology are illustrated in the following section.

2. Inflection of nouns and adjectives in thematic morphology

2.1. Thematic morphology: an introduction

The analysis I propose is developed within the so-called "thematic" approach to morphology. This label refers to a series of studies on inflection and word formation, realized mainly in France, within a word-and-paradigm framework (*cf.* Bonami & Boyé, 2007; Plénat, 2008, 2009). This model is fairly similar to the approach developed independently by Pirrelli & Battista (2000). In this approach the focus is more on the global structure of a lexeme's paradigm than on local relations between its forms. As observed above, having a unique underlying form is neither an issue, nor a specific property of regular *vs* irregular lexemes. The thematic approach is based on the observation that allomorphy and suppletion do not distribute randomly across lexemes, but follow systematic patterns. In fact, there are portions of the paradigm of a lexeme that systematically covary. Note that these subparadigms (which I will call "partition classes", as suggested by Pirrelli & Battista, 2000, or simply "stems", as explained in 1.1.3) can only be identified on the basis of this purely morphological property; in some cases, in fact, the forms belonging to the same partition class do not constitute a natural class, either phonologically or semantically. These subparadigms roughly correspond to Aronoff's (1994) morphomes. The identification of such a level of organization has proven useful to account for the distribution of several phenomena, such as syncretism (Corbett, 2007), defectiveness (Boyé & Cabredo-Hofherr, 2008), or diachronic change (Maiden, 1992).

The main assumptions on which thematic morphology is based may be resumed as follows:

- as observed above, only a subset of all the possible ways in which irregularity (allomorphy) may be distributed within a lexeme are attested. In other words, for a specific category in a given language it is possible to define a pattern for the distribution of allomorphy (which has been labelled its "stem space", Bonami & Boyé, 2003) specifying the maximal complexity that a lexeme of this category can reach. All lexemes displaying some irregularity have a pattern which is a variation of the stem space for their category.
- partition classes correspond to sets of forms in systematic covariation. Moreover, some partition classes covary more often than others, and should be considered, therefore, as more tightly connected (see the examples given below in 2.2.). It is then possible to propose a network of dependencies for the partition classes of a lexeme;
- for a regular lexeme the partition classes are connected by predictable functions. The only distinction which is made is between general phonotactic rules, which operate automatically, and morphological functions. These functions may correspond to identity or to a phonological rule, which may be more or less motivated, but no distinction is made between these types. For instance, for the functions linking the masculine and the feminine stem of French adjectives no distinction is made between identity (*e.g. pur – pure*), a phonological rule like (de)nasalization (*e.g. fin – fine*), or apparently unmotivated phonological relations (*e.g. beau – belle*). Note that the relations between the stems are non-oriented. This is coherent with the idea that lexemes do not have a unique base, and accounts for the fact that speakers are able to reconstruct paradigms, or a part of them, from any of their forms. For irregular lexemes, there are at least two stems that are not connected by a predictable function, and at least two forms have to be stored in the lexicon. This view of the lexemes allows us to have an explicit definition of what is a regular lexeme, and to consider regularity as a gradient property: a fully regular lexeme is a lexeme for which all the stems are linked by predictable functions and which can fully serve as a model for the inflection of other lexemes.

2.2. An illustration: French adjectives

As a first illustration of thematic morphology, I take the example of French adjectives. As in other Romance languages, the inflectional paradigm of French adjectives has four cells, corresponding to the combination of the values [masculine], [feminine] for the feature Gender and [singular], [plural]

for the feature Number. As observed above, for the majority of adjectives all these forms are homophonous, while for others the masculine is distinct from the feminine, and both forms have to be stored in the lexicon, since the relation between them does not always correspond to a synchronically active phonological rule. The adjectives PUR and BEAU illustrate respectively the first and the second case:

(5) a. PUR

	Masc	Fem
Sg	pyr	pyr
Pl	pyr	pyr

b. BEAU

	Masc	Fem
Sg	bo	bɛl
Pl	bo	bɛl

From the representation in (5) we can draw the following generalization, where A corresponds to the stem used for the masculine and B to the stem used for the feminine:

(6) a. PUR

	Masc	Fem
Sg	A	B
Pl	A	B

Stem A and stem B may or may not be stored independently in the lexicon. They are for BEAU, while for PUR we may postulate a 'rule' stating that the two are identical (B=A), which corresponds, as we have seen, to the default relation for French adjectives. Inflected forms too are linked to the stem corresponding to their cell by a function, in this case also identity:

(7) Masc Sg = A
Masc Pl = A
Fem Sg = B
Fem Pl = B

French adjectives, thus, have a two-stem stem space, and the stem space of regular lexemes (those like PUR) is a reduced version of the stem space of irregulars (like BEAU).[10]

10. In fact, the situation is slightly more complicated: a full model should take into account adjectives like SPÉCIAL ('special'), which have a distinct form in the masculine plural (*special* [spesjal] / *spéciaux* [spesjo]), and adjectives like GROS ('big'), which have a special masculine singular liaison form (*cf. un gro[z] avion* 'a big airplane') (*cf.* Bonami & Boyé,

Remember that regularity is considered to be a gradient property. Thus, it is quite easy to define PUR as a regular lexeme, since it exemplifies the relation between the two inflectional stems of an adjective which is the most common (more than 80%, *cf.* Table 2 above). We can define the adjectives illustrated in Table 2 as having a decreasing degree of regularity, since the possibility that they function as models for the inflection of other lexemes decreases with their frequency. Under this conception, irregularity may be linked to two explicit criteria: (i) the number of lexemes exemplifying a pattern and (ii) the fact that a pattern is expressed by a motivated or unmotivated phonological function. The alternations observed in LENT or SAIN (*cf.* Table 2), which are based on a simple phonological rule (consonant adjunction and (de)nasalization) are thus more regular than the one observed in BEAU, which is phonologically unmotivated ([o] ~ [ɛl]) and represented by a small number of lexemes (12 in the Lexique.org database). The real irregulars are those lexemes, which display a unique alternation, such as GENTIL ('kind', [i] ~ [ij]), SAOÛL ('drunk', [u] ~ [ul]), LAID ('ugly', [ɛ] ~ [ɛd]), etc. We may thus represent the "rule" linking the masculine (A) and the feminine (B) stem of French adjectives as a set of hierarchically ordered constraints on their form (*cf.* Boyé, to appear for a similar proposal concerning verbal inflection; = indicates identity):

(8) $= > \tilde{a} \sim \tilde{a}t > \emptyset \sim œz > \tilde{\varepsilon} \sim \varepsilon n > \ldots > i \sim ij \ldots$

So far, I have illustrated a simple example of the treatment that a thematic approach can offer for inflectional paradigms. However, the model is intended to deal with both inflectional and derivational phenomena. The notion of stem as illustrated above has proven to be useful for explaining some phenomena of allomorphy in derivation (see *e.g.* Bonami *et al.*, 2009 for an application to French deverbals; an explanation of this kind could also be imagined for the deverbals apparently constructed on the basis of a past participle illustrated in (2)). Plénat (2008: 1615-1616) analyzed some deadjectival nouns in *-ité*, and proposed that this suffix (as well as other suffixes, such as *-isme*, *-itude* or *-iser*) is attached to a specific stem of the adjective, which he called the "L stem".[11] As for inflection, this stem may be linked with other stems in the paradigm by identity (9a), by a phonologically motivated relation (9b) or by a fully unpredictable one (9c):

2005 for details). It should be remembered, moreover, that all plurals also have a liaison form ending in [z] (*cf. les belle[z] histories* 'the nice stories').

11. For "learned", as the majority of these derived forms are of Greek or Latin origin.

(9)

	A (masc.)	B (fem.)		L (*-ité*)
a. docile / docilité ‘obedient’	dɔsil	dɔsil	=	dɔsil
fécond / fécondité ‘fecund’	fekɔ̃	fekɔ̃d	=	fekɔ̃d
b. inférieur / infériorité ‘inferior’	ɛ̃ferjœr	ɛ̃ferjœr	œ ~ ɔ	ɛ̃ferjɔr
c. aigu / acuité ‘acute’	ɛgy	ɛgy	≠	aky
rond / rotondité ‘round’	rɔ̃	rɔ̃d	≠	rɔtɔ̃d

Adjectives such as FÉCOND or ROND[12] must specify both their inflectional stems in the lexicon. However, the data in (9) show another interesting fact, and namely that the L stem, when it is not fully suppletive, is always directly connected with the B stem, even when the latter is not itself linked to the A stem. In other words, if the L stem is identical to the A stem, the B stem is necessarily identical too. Consequently, we can propose a spatial representation of the dependency between these three stems, in which the B stem is directly connected with both the A and the L stem, while the latter are not directly connected to each other: a speaker can only go from the A stem to the L stem via the B stem (I indicate identity as the default form):

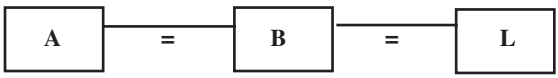


Figure 1: (partial) network of dependencies for French adjectives

Figure 1 represents the network of dependencies for a fully regular adjective, *i.e*. an adjective for which all the stems are connected by a predictable function. As observed above, for (partially) irregular lexemes some of the

12. An anonymous reviewer correctly pointed out that the adjective ROTOND also exists in French, and that it could be the “right” base of *rotondité*. The question is if it is legitimate and necessary to assume that derivational rules have one “right” base. It is certainly true from an etymological point of view that *rotondité* is derived from *rotond*. However, while the former does not belong to the standard vocabulary of French (the *Trésor de la langue française* labels it as *vieilli ou littér*., old-fashioned or literary), ROND is the standard lexeme used to express the meaning ‘round’. Consequently, a connection may be made in the lexicon, at least for some speakers, between it and the quality noun *rotondité*.

connections are interrupted, and the content of each cell must be specified in the lexicon, as illustrated below for the adjectives FÉCOND, AIGU and ROND:

Figure 2a: network of dependencies for FÉCOND

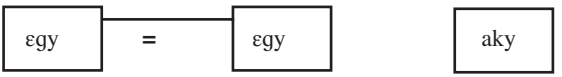


Figure 2b: network of dependencies for AIGU

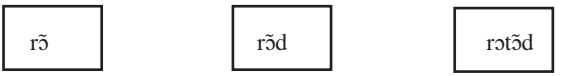


Figure 2c: network of dependencies for ROND

3. The representation of lexemes in inflectional languages

So far, I have defended the idea that inflectional forms or sets of inflectional forms (stems) need not be decomposed, and can be considered as stored in the lexicon as global forms, with morphological competence specifying the relations between them. I have illustrated this fact with French examples, which are quite straightforward in this respect. Modern French inflected forms, especially for adjectives, are the outcome of phonological changes that applied to earlier forms containing overt suffixes. Although these affixes have disappeared, some of their phonological effects are still visible. Consequently, the inflected forms of French adjectives are generally recognized as synchronically indecomposable: nobody has seriously proposed, to my knowledge, to consider, for instance, that [d] is a morpheme with the meaning [feminine] in *féconde* or *ronde*. However, I claim that even in languages which apparently have overt inflectional markers the non-decomposition of forms provides a better and more insightful analysis. I start my demonstration by briefly considering a simpler example (Catalan), before turning to a more complicated one (Italian).

3.1. Inflection of Catalan nouns and adjectives[13]

Catalan nouns and adjectives apparently display overt affixes for gender and number inflection. The singular number is unmarked, while the plural number is systematically marked by [s]. The masculine gender may be unmarked, but there are at least some masculine forms ending in [u] (graphically *-o*); the feminine is generally marked by [ə] (graphically *-a*). According to traditional morphemic analyses (*cf.* Clua, 2002 for an overview), [s] is the plural suffix, [u] is the masculine suffix for some nouns and adjectives, and [ə] is the feminine suffix. One of the arguments used to affirm that these elements are inflectional morphemes is that they never surface in derivation. This analysis encounters at least two problems: (i) there are some masculine nouns ending in [ə] (graphically *-e*) and some feminine nouns either unmarked or ending in [u] (10a); (ii) the vowels in question are not erased in all derived forms, while there are other sequences which are not considered to be inflectional markers that do not surface in derived forms (10b):

(10) a. cotxe$_{\text{MASC}}$ 'car'
sal$_{\text{FEM}}$ 'salt'
moto$_{\text{FEM}}$ 'motorcycle'
b. ritu → ritual 'rite'/'ritual'
virus$_{\text{MASC}}$ → viral 'virus'/'viral'
bilis$_{\text{FEM}}$ → biliar 'bile'/'biliary'

The same alternation between an unmarked form (11a) or a form marked in [u] (11b) for the masculine and a form marked in [ə] for the feminine is observed with adjectives. In this case too, however, there are some adjectives whose feminine is unmarked (11c), and others whose masculine ends in [ə] (11d) (in this case, the distinction between masculine and feminine is purely graphic):

(11) a. prim prima 'thin'
b. flonjo flonja 'soft'
c. gran gran 'big'
d. pobre pobra 'poor'

For adjectives of the type (11d) a phonological explanation is often invoked: the [ə] is, in these cases, an epenthetic vowel that resolves a sequence otherwise impossible word-finally. However, while it is the case for POBRE, other adjectives displaying a final sequence, which would be acceptable in final position also end in [ə]: *còmode* ('comfortable'), *belga* ('Belgian').

As in French, we can consider that the stem space of Catalan adjectives includes (at least) two stems, one for the masculine forms and one for the

13. The analysis of Catalan I propose is the result of joint work with Aurélie Guerrero, to whom I am grateful.

feminine forms.[14] For some adjectives, in fact, these forms display an alternation that cannot be accounted for simply by phonological rules (*cf.* the data in (12)):

(12)	Masc Sg	Masc Pl	Fem Sg	Fem Pl	
	clar ['kla]	clars ['klas]	clara ['klarə]	clares ['klarəs]	'clear'
	car ['kar]	cars ['kars]	cara ['karə]	cares ['karəs]	'dear'

If this is the case, there is no obstacle for considering that the phoneme [ə] is part of the feminine stem, and that, in default cases, at least for adjectives like PRIM and CAR, the masculine (A) and the feminine (B) stems are linked by a function of the type X – Xə. For adjectives like CLAR, on the other hand, both stems have to be stored in the lexical representation of the lexeme. Of course, this function only holds for one class of adjectives. For the other types (represented in (11) by FLONJO and by GRAN and POBRE, respectively), other default functions are active, namely Xu – Xə for the first type, and X – X (identity) for the second. Thus, the final vowels are disconnected from the expression of gender, a desirable effect, since, as we have seen, they may or may not appear in forms realizing a specific gender. The three types of adjectives we have identified greatly resemble the inflectional classes of Italian illustrated in (1). In fact, to sum up, what we have in Catalan is an inflectional stem space common to all adjectives (13a), three inflectional classes, each specifying a function linking the two stems of the stem space (13b), and a series of rules for the construction of actual inflected forms (13c):

(13) a.

	Masc	Fem
Sg	A	B
Pl	A	B

b. Class I: X – Xə (B = A+ə)
Class II: Xu – Xə (B = A – u+ə)
Class III: X – X (B = A)

c. Masc Sg = A
Masc Pl = A+s
Fem Sg = B
Fem Pl = B+s

14. Once again, this is not the full story: the stem space is certainly a bit more complicated, since some adjectives display a variation between the masculine singular and plural, *cf. sord* ['sort] – *sords* ['sors] ('deaf').

What is stored in the lexical representation of Catalan adjectives (and the same holds for nouns) is then a fully specified word form and a pattern of relations between the stems. In the default case, one form is sufficient, and the relations in (13) permit the reconstruction of the whole paradigm. The final vowels have no inflectional status (they do not directly mark gender), and they are not theme vowels, since they do not indicate inflectional class membership, which is ambiguous (a consonant-final masculine may belong either to Class I or to Class III, and a feminine in [ə] to any of the three classes). Class membership is in fact indicated by the relation pattern attached to the lexical representation and there is no difference, in substance, between classes in this sense and the inflectional patterns identified for French adjectives above.

3.2. Inflection of Italian adjectives and nouns

Italian represents a more complicated example than Catalan, since in Italian all inflected forms are typically marked by a specific vowel. Italian adjectival classes are illustrated in (1). The adjectival system of Italian is a conflation of the nominal one, which is normally regarded as including at least five classes (*cf.* D'Achille & Thornton, 2003; Acquaviva, 2009):[15]

(14)

class	sg./pl.	gender	example
I	*-o*/*-i*	Masc	tavolo/i ('table')
II	*-a*/*-e*	Fem	casa/e ('house')
III	*-e*/*-i*	Masc/Fem	fiore/i ('flower')
IV	*-a*/*-i*	Masc	pilota/i ('pilot')
V	invariable	Masc/Fem	specie ('species')
		kiwi ('kiwi')	
		città ('city')	
		bar ('bar')	

Apart from Class V, final vowels seem to have a morphemic status in Italian, since they systematically appear in inflectional forms realizing specific morphosyntactic properties. An argument which is often put forth in support of this hypothesis is the fact that the final vowel never surfaces in derived forms, especially before a suffix beginning with a vowel (*cf. tavolo* → *tavolino* 'small table'). On the basis of these arguments, some authors consider that Italian (nominal) morphology is exclusively morpheme-based (*cf.* Peperkamp,

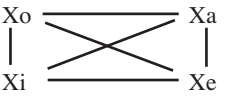
15. The column "gender" in (14) indicates the prevailing gender for each class, although there are some exceptions (*e.g.* feminine *mano*/*i* 'hand' in Class I). As shown, Class V contains invariable nouns, which can end in a stressed vowel (*città*), in an unstressed vowel (*specie*, *kiwi*), or in a consonant (*bar*). Another small group of nouns has a masculine singular in *-o* and a feminine plural in *-a* (*cf. uovo*/*a* 'egg'), but it is not certain that it constitutes a "real" inflectional class (*cf.* Acquaviva, 2008: 123-161 for details).

1995; Crocco Galèas, 1998). Under this hypothesis, lexical morphemes do not carry any final vowel (which are introduced in syntactic constructions), and class membership is specified by a diacritic attached to each base. Another hypothesis, in a word-based framework, has been proposed by Scalise (1983; 1984). According to Scalise, lexemes are stored in the lexicon in their full form, including a final vowel, which is a theme vowel indicating class membership. A rule of final vowel erasure operates in derivation, which is responsible for the final form of derived words. This hypothesis is preferable, since final vowel erasure is a general phenomenon in Italian. It can be shown, in fact, that it operates also on clearly monomorphemic words ending in a vowel, such as proper names, borrowings, adverbs, numerals, etc.:

(15)	Milano	→	milanese	'Milan'/'Milanese'
	koala	→	koalino	'koala'/'small koala'
	bene	→	benissimo	'well'/'very well'
	venti	→	ventesimo	'twenty'/'twentieth'

In Montermini (2003) I claimed that the rule in question is a morphophonological rule, which operates under the influence of various factors, including the similarity between the two vowels in contact, the nature of the vowel potentially targeted by erasure ([i] and [u] being the most resistant to erasure), etc. However, Scalise's hypothesis has a very serious drawback, since it considers that the stem of the adjective or of the noun that is stored in the lexicon is systematically homophonous with its citation form (the singular for nouns, and the masculine singular for adjectives), clearly not a theoretically relevant object. In fact, nothing indicates which form of the lexeme is actually stored in a speaker's lexicon. It is not rare, for instance, that Italian children produce such forms as *caria* and *cecio*, instead of the expected *carie* ('tooth decay') and *cece* ('chick-pea') (these are actual productions of my own children). Clearly, in these cases, what they have stored in their lexicon is the plural form of these lexemes, which is more frequent, from which they reconstruct the singular on the basis of the most common patterns (respectively *-a*/*-e* for feminines and *-o*/*-i* for masculines. Thus, in Italian, as in Catalan, for regular lexemes the memorization of one full form is sufficient to inflect the whole lexeme. However, nothing indicates exactly which form is memorized. The situation of Italian inflectional classes is slightly more complicated than that of Catalan and of French, since in Italian for nouns and adjectives there is not a basic form from which the others may be obtained by identity or by adjunction of phonological material. Rather, the inflectional patterns for Italian nominals constitute a network, not a list of functions. Let us consider, for instance, the main adjectival inflectional classes illustrated in (1), the ones represented by BELLO and FORTE, respectively. A complete representation of their inflectional patterns would be as follows:

(16) a. BELLO



b. FORTE

Under this analysis, the most specified form that can be proposed for the lexical representation of these adjectives is XV (*cf.* (4) above), and the same holds for nouns (*cf.* Passino, 2009 that proposes, in a different framework, an analysis in which nominal stems end in an empty nucleus). This representation is a shortcut to express the fact that it is sufficient that any of the inflected forms of the lexeme, along with the relevant inflectional pattern, is memorized in the lexicon in order for that lexeme to be correctly inflected. As for the other languages, the different patterns for the inflection of a specific category of lexemes may be considerer as hierarchically ordered on the basis of their frequency, and thus of their capacity to serve as models for the inflection of other lexemes. The relative frequency of adjectival inflectional classes was given in Table 1 above (for similar statistics on Italian nouns *cf.* D'Achille & Thornton, 2003: 213). For instance, borrowings may be assigned to an inflectional class (and to the default gender for that class) on the basis of their phonological form, and in particular of their final vowel (*cf. tapiro*/*i* 'tapir', *macaco*/*i* 'macaque', *kimono*/*i vs balalaika*/*e*, *corrida*/*e*, *sauna*/*e*, more examples in Thornton, 2003). Of course, this frequency-based hierarchy interacts with other parameters, such as morphological (*e.g.* the presence of a suffix or of a suffix-like sequence) and semantic criteria for the assignment of an unknown novel word to a specific class (*cf.* Thornton, 2001, 2003 for an analysis of class assignment in Italian).

A last observation concerns the final vowel of Italian nouns and adjectives. In what precedes I have suggested that this vowel has no specific morphemic value (it does not mark gender or number), nor it is a theme vowel, that would specify, for instance, inflectional class membership. However, it cannot be denied that final vowels have a special role in Italian morphology and phonology. It is a fact, for instance, that the majority of native words end in a vowel, including demonstratives (*questo* 'this'), quantifiers (*tutto* 'all'), and uninflected words, such as adverbs (*domani* 'tomorrow', *prima* 'before'). It is not by chance that, as noted above, the final vowel of an originally monomorphemic word (such as a borrowing, a clipping and even a proper noun) is often reinterpreted as an inflection marker:

(17)

la sauna	le saune	'sauna(s)'
la para	le pare	(< paranoia (id.))
Roma	le due Rome[16]	'Rome'/'the two Romes'

16. More than 50,000 hits for "le due Rome" on Google (search of April 2010).

A possible characterization of these vowels, which is compatible with the emergent model based on analogy I adopt, is to consider that they are "phonotactic markers", as apparently suggested by Acquaviva (2009: 55) on the basis of the proposal made by Harris (1996) for Spanish. These markers guarantee that a word is a well-formed phonological word. In the most prototypical Italian lexemes they also guarantee that a word is morphologically well-formed.

4. Conclusion

In this paper I have proposed an analysis of nominal and adjectival inflection in some Romance languages within a word (lexeme)-based model of morphology. The approach I propose is emergent, in that it considers that morphological "rules" emerge as patterns on the basis of generalizations speakers make about the existing lexicon. Some basic assumptions of this approach are:

- the lexicon does not necessarily store information non-redundantly. Rather, a model in which full word forms are stored and stand in a network of relations better accounts for the behaviour of speakers when they extend existing patterns to novel / unknown words; morphology is intended to account for the units that "must" be stored in the lexicon and the units that do not necessarily need to, but "can" be stored;
- frequency is an important factor in the structuring of morphological competence. The most frequent patterns tend to be identified as rules. However the distinction between more regular and less regular lexemes is a quantitative and not a qualitative one;
- the existence of a unique underlying form is not a theoretical necessity, nor a defining property of regular *vs* irregular lexemes;
- the decomposition of complex (inflected or constructed) words is a theoretical or descriptive desideratum. However, it is not certain that the units identified correspond to units that have a real theoretical status in a speaker's morphological competence.

Concerning the last point, it should be observed that recent developments within symbolic approaches, such as OT, recognize the fact that full paradigms (or subparadigms) may be the input for morphophonological rules (*cf.* McCarthy, 2005), and admit that the lexicon may store stems as full forms (*cf.* in particular Bermúdez Otero, 2009). The analysis proposed here could therefore be implemented in a model of this type.

REFERENCES

Acquaviva, Paolo (2008). *Lexical Plurals. A Morphosemantic Approach*. Oxford: Oxford University Press.

Acquaviva, Paolo (2009). The structure of the Italian declension system. In Montermini. F., Boyé, G. & Tseng, J. (eds.), *Selected Proceedings of the 6th Décembrettes. Morphology in Bordeaux*: 50-62. Somerville, Ma.: Cascadilla Press.

Albright, Adam (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language* 78: 684-709.

Albright, Adam (2009). Modelling analogy as probabilistic grammar. In Blevins, J.P. & Blevins, J. (eds.), *Analogy in Grammar*: 185-213. Oxford: Oxford University Press.

Aronoff, Mark (1994). *Morphology by Itself*. Cambridge, Ma: MIT Press.

Baayen, R. Harald (2003). Probabilistic approaches to morphology. In Bod, R., Hay, J. & Jannedy S. (eds.), *Probabilistic Linguistics*: 229-87. Cambridge, Ma: MIT Press.

Baayen, R. Harald; McQueen, James M.; Dijkstra, Ton; Schreuder, Robert (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen & Schreuder (eds.), *Morphological Structure in Language Processing*: 355-90. Berlin: Mouton de Gruyter.

Bermúdez Otero, Ricardo (2009). The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. Manuscript. University of Manchester.

Blevins, James P. (2006). Word-based morphology. *Journal of Linguistics* 42: 531-573.

Blevins, James P.; Blevins, Juliette (2009). Introduction: Analogy in grammar. In Blevins & Blevins (eds.), *Analogy in Grammar*: 1-12. Oxford: Oxford University Press.

Bonami, Olivier; Boyé, Gilles (2003). Supplétion et classes flexionnelles dans la conjugaison du français. *Langages* 152: 102-26.

Bonami, Olivier; Boyé, Gilles (2005). Construire le paradigme d'un adjectif. *Recherches Linguistiques de Vincennes* 34: 77-98.

Bonami, Olivier; Boyé, Gilles (2007). Remarques sur les bases de la conjugaison. In Delais-Roussarie & Labrune (eds.), *Des sons et des sens. Données et modèles en phonologie et en morphologie*: 77-90. Paris: Hermès-Lavoisier.

Bonami, Olivier; Boyé, Gilles; Kerleroux, Françoise (2009). L'allomorphie radicale et la relation flexion-construction. In Fradin, Kerleroux & Plénat (eds.), *Aperçus de morphologie du français*: 103-25. Saint-Denis: Presses Universitaires de Vincennes.

Boyé, Gilles (à paraître). Régularité et classes flexionnelles dans la conjugaison du français. In Roché, Boyé, Hathout, Lignon & Plénat, *Des unités morphologiques au lexique*.

Boyé, Gilles; Cabredo Hofherr, Patricia (2008). Defectivity as stem suppletion in French and Spanish verbs. Paper presented at the workshop *Defective Paradigms: Missing Forms and What They Tell Us*. London, 10-11 April 2008.

Burzio, Luigi (2002). Surface-to-surface morphology: when your representations turn into constraints. In Boucher (ed.), *Many Morphologies*: 142-77. Somerville, Ma: Cascadilla Press.

Burzio, Luigi (2003). Output-to-output faithfulness: The Italian connection. *Lingue e linguaggio* 2/1: 69-104.

Bybee, Joan L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

BYBEE, Joan L. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

CLUA, Esteve (2002). Gènere i nombre en els noms i en els adjectius. In Solà, Lloret, Mascaró & Pérez Saldanya (eds.), *Gramàtica del català contemporani*: 485-534. Barcelona: Editorial Empúries.

CORBETT, Greville G. (2007). Deponency, syncretism, and what lies between. In Baerman, Corbett, Brown & Hippisley (eds.), *Deponency and Morphological Mismatches*: 21-43. Oxford: Oxford University Press.

CROCCO GALÈAS, Grazia (1998). La base dei processi morfologici in italiano. *Studi di grammatica italiana* XVII: 245-72.

D'ACHILLE, Paolo; THORNTON, Anna M. (2003). La flessione del nome dall'italiano antico all'italiano contemporaneo. In Maraschio & Poggi Salani (eds.), *Italia linguistica anno Mille – Italia linguistica anno Duemila. Atti del XXXIV Congresso internazionale della Società di Linguistica Italiana*: 211-30. Roma: Bulzoni.

DE MAURO, Tullio; MANCINI, Federico; VEDOVELLI, Massimo; VOGHERA, Miriam (1993). *LIP. Lessico di frequenza dell'italiano parlato*. Milano: Etas libri.

DERWING, Bruce L.; SKOUSEN, Royal (1988). Morphology in the mental lexicon: A new look at analogy. In Booij & van Marle (eds.), *Yearbook of Morphology 1989*: 56-71. Dordrecht: Kluwer.

DIXON, Robert M.W.; AIKHENVALD, Alexandra Y. (2002). Word: a typological framework. In Dixon & Aikhenvald (eds.), *Word: A Cross-Linguistic Typology*: 1-41. Cambridge: Cambridge University Press.

FRADIN, Bernard (2003). *Nouvelles approches en morphologie*. Paris: PUF.

FRADIN, Bernard; MONTERMINI, Fabio; PLÉNAT, Marc (2009). Morphologie grammaticale et extragrammaticale. In Fradin, Kerleroux & Plénat (eds.), *Aperçus de morphologie du français*: 21-45. Saint-Denis: Presses Universitaires de Vincennes.

HALLE, Morris (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry* 4/1: 3-16.

HARRIS, James (1996). The syntax and morphology of class marker suppression in Spanish. In Zagona (ed.), *Grammatical Theory and Romance Languages*: 99-122. Amsterdam: John Benjamins.

JACKENDOFF, Ray (1975). Morphological and semantic regularities in the lexicon. *Language* 51/3: 639-71.

JACKENDOFF, Ray (2002). *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

MAIDEN, Martin (1992). Irregularity as a determinant of morphological change. *Journal of Linguistics* 28: 285-312.

MATTHEWS, Peter H. (1972). *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.

MCCARTHY, John J. (2005). Optimal paradigms. In Downing, Hall & Raffelsiefen (eds.), *Paradigms in Phonological Theory*: 170-210. Oxford: Oxford University Press.

MONTERMINI, Fabio (2003). Appunti sulla cancellazione di vocale in derivazione. In Bisetto, Iacobini & Thornton (eds.), *Scritti di morfologia in onore di Sergio Scalise in occasione del suo 60° compleanno*: 171-88. Cesena: Caissa Italia.

MONTERMINI, Fabio (2006). La représentation phonologique des unités lexicales et la question de l'allomorphie. Sur quelques affixes problématiques de l'italien. *Cahiers de grammaire* 30: 291-305.

Passino, Diana (2009). An element-based analysis of Italian nominal inflection. In Montermini, Boyé & Tseng (eds.), *Selected Proceedings of the 6th Décembrettes. Morphology in Bordeaux*: 63-75. Somerville, Ma: Cascadilla Press.

Peperkamp, Sharon (1995). Prosodic constraints in the derivational morphology of Italian. In Booij & van Marle (eds.), *Yearbook of Morphology 1994*: 207-44. Dordrecht: Kluwer.

Pinker Steven; Ullman Michael T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences* 6.11: 456-63.

Pirrelli, Vito; Battista, Marco (2000). The paradigmatic dimension of stem allomorphy in Italian. *Rivista di linguistica* 12/2: 307-80.

Plag, Ingo (2009). Morphological complexity: Inflection classes and probabilistic allomorph selection. Paper presented at the ESF Exploratory Workshop *Words in Action*. Pisa, 11-14 October 2009.

Plénat, Marc (2008). Le thème L de l'adjectif et du nom. In Durand, Habert & Laks (eds.), *Congrès Mondial de Linguistique Française CMLF '08* : 1613-26. Paris: Institut de Linguistique Française.

Plénat, Marc (2009). Le conditionnement de l'allomorphie lexicale en français. *Mémoires de la Société de Linguistique de Paris*. Nouvelle série XVII: 119-40.

Scalise, Sergio (1983). *Morfologia lessicale*. Padova: Clesp.

Scalise, Sergio (1984). *Generative Morphology*. Dordrecht: Foris.

Stemberger Joseph P.; MacWhinney Brian (1986). Frequency and lexical storage of regularly inflected forms. *Memory and Cognition* 14.1: 17-26.

Stump, Gregory T. (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Thornton, Anna M. (1990/1991). Sui deverbali italiani in *-mento* e *–zione*. *Archivio glottologico italiano* LXXV.2: 169-207 / LXXVI.1: 79-102.

Thornton, Anna M. (2001). Some reflections on gender and inflectional class assignment in Italian. In Schaner-Wolles, Rennison & Neubarth (eds.), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler on the Occasion of his 60th Birthday*: 479-87. Torino: Rosenberg & Sellier.

Thornton, Anna M. (2003). L'assegnazione del genere in italiano. In Sánchez Miret (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*. Vol. I: 467-81. Tübingen: Niemeyer.

Tranel, Bernard (1981). *Concreteness in Generative Phonology. Evidence from French*. Berkeley: University of California Press.

Xu, Zheng (2007). *Inflectional Morphology in Optimality Theory*. Doctoral dissertation. Stony Brook University.

RÉSUMÉ

Cet article traite de la manière dont les unités lexicales sont stockées dans la mémoire des locuteurs, de la manière dont la flexion et la dérivation interagissent, et de la définition de quelques notions morphologiques, comme thème, racine, mot-forme, etc. Plus largement, deux modèles concurrents de la compétence morphologique y sont discutés : un modèle à « unités plus règles » et un modèle global basé sur l'analogie. L'analyse est conduite en particulier sur trois langues romanes : français, catalan et italien. Les données sont analysées dans le cadre d'un modèle de morphologie basé sur les mots. Les mots-formes sont considérés comme les unités de base de la mémorisation lexicale d'un point de vue universel. Ces formes sont organisées en des unités plus larges et plus abstraites, les lexèmes, qui correspondent *grosso modo* aux entrées d'un dictionnaire. Il est soutenu que les autres unités traditionnellement identifiées dans l'analyse morphologique, comme les thèmes, les racines, etc., peuvent fonctionner comme des unités d'organisation dans certains cas, mais ne doivent pas être considérées comme des unités morphologiques sur une base universelle.

MOTS-CLÉS

Morphologie, lexique, allomorphie, paradigmes, italien, français, catalan.