

Chapter 13

Troubles with flexemes

Anna M. Thornton

University of L'Aquila

This paper investigates an aspect of the notion *flexeme* (French *flexème*), introduced by Fradin & Kerleroux (2003), Fradin (2003). After a brief review of how this concept developed in these authors' work, and of how these authors conceive of lexemes (Section 2), the relation between flexemes and overabundance (Thornton 2011, 2012) is explored. Overabundance is introduced in Section 3, and Section 4 is devoted to some case studies, from Italian and other languages. It is shown that a single lexeme can map to more than one flexeme – and overabundance results from this mapping. Besides, it is shown that flexemes differing from each other in parallel ways can have various relations with lexemes: in some cases, mapping to different flexemes distinguishes two lexemes that are homophonous in their citation form (e.g., Italian *SUCCEDERE*¹ 'happen' with PST.PTCP *successo* and *SUCCEDERE*² 'succeed' with PST.PTCP *succeduto*), while in other cases flexemes that differ from each other in a way parallel to the previous one map to a single overabundant lexeme (e.g., Italian *PERDERE* 'lose' with PST.PTCP *perso* and *perduto*). I conclude that the distinction between lexemes and flexemes first proposed by Fradin & Kerleroux (2003) and Fradin (2003), as well as their definition of lexeme, based on semantic and constructional coherence rather than on inflectional coherence, is useful even beyond the area of lexeme formation for which it was originally proposed.

1 Introduction

In a paper titled "Troubles with lexemes", Bernard Fradin and Françoise Kerleroux (2003) laid the bases for a critique of the commonly held notion of lexeme, drawing data from the realm of word-formation. They observed at the beginning of their paper:

the lexeme is supposed to constitute one lexical unit. **This unicity is guaranteed by inflection on the one hand** and by the semantic content of the lexeme, which is supposed to be unique, on the other (Fradin & Kerleroux 2003: 177, emphasis mine).

They proceeded then to show that the objects to which word-formation rules apply – which they propose to call lexemes, partially modifying the usual definition of this term – are semantically fully specified objects, that are, however, unspecified for inflection. In



the concluding section of that paper, they propose to distinguish three different theoretical entities: lexemes (“lexical individuals defined by the conjunction of three properties: category, underspecification for inflection, full specification for meaning”, Fradin & Kerleroux 2003: 193), syntactic words (which are inflected, categorized, and fully specified for meaning), and a third entity, which they propose to call *inflecteme* in English and *flexème* in French (see also Fradin 2003: 259). Objects of this third type are categorized, uninflected and underspecified for meaning.

In this short contribution, I will discuss some aspects of these entities that have come to the fore of the debate in morphology after the publication of Fradin & Kerleroux (2003) and Fradin (2003). I prefer to refer to these units as *flexèmes*, because I think that the intentional and witty phonological and orthographic overlap with *lexème* ‘lexeme’ is too good to be lost, and as an *hommage* to the authors who first proposed this term. Following Fradin (forthcoming), in this paper I will use the English adaptation *flexeme*.

The paper is organized as follows: Section 2 reviews the development of the concept of flexeme; Section 3 introduces the concept of overabundance in inflectional paradigms; Section 4 presents several case studies from Italian and other languages, illustrating cases in which a single lexeme is overabundant in one or more cells, i.e., maps to two distinct flexemes; Section 5 concludes.

2 What are flexemes?

In different contributions by Bernard Fradin, sometimes in collaboration with Françoise Kerleroux, the concept of *flexème*/flexeme is presented differently: its coverage seems to have grown with time, probably in consequence of our growing understanding of the workings of inflectional morphology in the early years of the third millennium.

In Fradin & Kerleroux (2003: 193) the concept seems to be equivalent to that of *stem* (in the sense, e.g., of Aronoff 1994):

This unit [i.e., the inflecteme/*flexème*] lacks semantic specification since it functions as the “inflectional stem”.

However, the authors seem to have something more than just a single stem in mind, since immediately after this definition they observe: “This is correlated to the fact that “no semantic constraints hangs [sic] over the application of inflectional rules” (Corbin 1987: 6)”. So the idea that flexemes have to do with instructions for building all the inflected forms that realize a lexeme seems to have been present already in Fradin & Kerleroux (2003).

Fradin (2003: 259) states that

Les flexèmes [...] comportent [...] des informations relevant [...] du syntactique interne (les différents thèmes flexionnels, sous forme hiérarchisée, s’il en existe plusieurs [...]).

So the concept of flexeme seems to have developed from being used to refer to a stem to being used to refer to the whole stem-set of a lexeme. In Fradin (forthcoming) a new development appears.¹ The author, dealing with verbs, distinguishes between verbs as morphological units, called “morphological verbs”, and verbs as lexical units, called “verbal lexemes”. He states that “[m]orphologically, a V is defined by its inflectional paradigm”, and maintains that the two French verbs *RESSORTIR*¹ ((de Y): *il ressort, il ressortait...*) ‘go out again’ and *RESSORTIR*² ((à Y): *il ressortit, il ressortissait...*) ‘come under’ “constitute distinct ‘flexemes’, see Fradin & Kerleroux (2003) [...] because the set of their word-forms is not identical” (Fradin forthcoming: 4).

In this passage, Fradin attributes to Fradin & Kerleroux (2003) a fully developed concept of flexeme, in which a flexeme contains all the information needed to generate all the inflectional forms in a paradigm: not only the information about which stem to select, but also inflectional class and realization rules for the different inflected forms. Roughly, it seems to me, a flexeme now corresponds to the entities called *form paradigm* and *realized paradigm* in paradigm-linkage theory (Stump 2016). Fradin (forthcoming) also equates the notion of flexeme with that of Paradigm Identifier adopted by Bonami & Tribout (2012). In turn, Bonami & Tribout (2012) state that their notion of Paradigm Identifier “[c]aptures Fradin & Kerleroux (2003)’s notion of a flexeme: a family of lexemes with the same inflectional paradigm” (Bonami & Tribout 2012: slide 16).²

Papers such as Fradin (forthcoming) and Bonami & Tribout (2012) address the question of how to deal with objects that are semantically different but morphologically identical, such as *CIRAGE*¹ ‘polishing’ and *CIRAGE*² ‘shoe polish’, or *PERLER*¹ ‘sew beads on’ and *PERLER*² ‘form beads on’, which share a flexeme (a form paradigm and a realized paradigm) but are different lexemes.³

In this paper, on the contrary, I will explore the issue of objects that are the same lexeme, in the sense of Fradin & Kerleroux (2003) and Fradin (2003, forthcoming), but can be realized, to variable degrees, by different flexemes.

3 Overabundance

In recent years, attention has been drawn to the phenomenon of overabundance in inflectional paradigms (Thornton 2011, Stump 2016: 147-151). Overabundance is defined as the situation in which two or more forms are available to realize the same cell in an inflectional paradigm; in terms of paradigm linkage theory, one content cell has more than one realization. Stump (2016: 148) gives an example from English. Consider the verbs *SEEM*, *MEAN*, and *DREAM*, and the realizations of their past tense: ⟨*SEEM*, {past}⟩ is realized by *seemed*, ⟨*MEAN*, {past}⟩ is realized by *meant*, and ⟨*DREAM*, {past}⟩ can be realized either by *dreamed* or by *dreamt*. The two (or more) forms that realize the same cell are sometimes called cell mates (Thornton 2011).

¹The notion of flexeme is not mentioned in Fradin & Kerleroux (2009).

²The notion of Paradigm Identifier is clearly articulated by Bonami & Crysmann (this volume).

³This phenomenon is labelled “homomorphy” by Stump (2016: 65): “homomorphic lexemes are lexically and semantically distinct but alike in every detail of their morphology”. English examples are *WEAR*¹ ‘have on (an article of clothing)’ and *WEAR*² ‘erode’.

How does overabundance relate to the notion of flexemes? Does the existence of distinct but synonymous realizations for a given content cell force us to recognize distinct flexemes linked to a single lexeme?

Fradin (forthcoming) analyzes cases such as *PERLER*¹ ‘sew beads on’ and *PERLER*² ‘form beads on’ as distinct lexemes linked to the same flexeme. The case of *dreamed* ‘dream.PST’ and *dreamt* ‘dream.PST’ appears to be a mirror image of this case, with distinct flexemes linked to a single lexeme. The existence of such a state of affairs would be predicted in Fradin’s theory, in which lexemes, defined as categorized and semantically fully specified but uninflected objects, are autonomous from the flexemes that provide instructions for the realization of their inflected forms. Recognizing the possibility that a single lexeme may be linked to two (or more) flexemes implies that a difference in inflectional realization cannot be invoked as one of the criteria that allow to distinguish between different lexemes vs. simply different senses/acceptations of a polysemous lexeme, as was sometimes done in traditional discussions of the homonymy/polysemy distinction (see e.g. Ullmann 1957: 127–132).⁴ Indeed, flexemes that are distinct in parallel ways may map to a single lexeme or to distinct lexemes – where the criterion for recognizing distinct lexemes is semantic and constructional difference, as proposed by Fradin & Kerleroux (2003, 2009) and Fradin (2003, forthcoming).

In the following section, I will review some data that show that the mapping between flexemes and lexemes can be of several kinds.

4 Non-canonical mappings between lexemes and flexemes

In this section, I will present data, mostly from well-studied cases in familiar languages, that show how one and the same difference in inflectional realization may map either to distinct lexemes or to a single overabundant lexeme.

4.1 Case study 1: Noun plurals

Nouns in which apparently more than one plural form pairs with a single singular form are very easy to find in language descriptions. Usually authors assume, at least implicitly, the admittedly vaguely defined criterion of ‘difference in meaning’ to decide whether specific cases represent distinct lexical items with homophonous singular forms or a single lexical item which is overabundant in its plural cell(s). Besides, since data are usually found in works which aim at description rather than at theoretical analysis, often authors leave the matter undecided, because it is not necessary for descriptive purposes to establish whether a certain case is an instance of homonymy or polysemy; on the other hand, cases in which no semantic distinction is observable between two or more different plural forms are usually highlighted by authors of descriptions.

⁴Remember also the observation by Fradin & Kerleroux (2003: 177) quoted in Section 1, that unicity of a lexeme “is guaranteed by inflection” as well as by the semantic content.

Cases such as the English and Breton ones in (1) and (2) are typical:

- (1) English (Aronoff 2000: 347)
- a. SG *brother* PL *brothers*
'male sibling'
 - b. SG *brother* PL *brethren*
'fellow member of a profession, society or sect'
- (2) Breton (Trépos 1980)⁵
- a. SG *eskob* PL *eskibien*
'bishop'
 - b. SG *eskob* PL *eskobou*
'kingpin'⁶

In these cases most authors argue that the meanings of the two items are sufficiently distinct to allow us to consider them as distinct lexemes, which happen to be homophonous in their singular form.⁷ In these cases, then, we have a 1:1 mapping between lexemes and flexemes, with the extra quirk represented by the fact that two distinct flexemes have homophonous singular forms.

However, by perusing the whole description of Breton noun plural offered by Trépos (1980), we discover that 'bishop' can have as many as three different plural forms (3a), and the same is true for 'coat' (3b):

- (3) Breton (Trépos 1980: § 149)
- a. SG *eskob* PL *eskibien/eskobed / eskeb*
'bishop'
 - b. SG *mantell* PL *mentell/mentellou/mentilli*
'coat'

A similar situation is common in Modern Standard Arabic, where nouns often have several plural forms; authors of descriptions usually comment on when they would prefer to assign the different plural forms to distinct lexemes, on the basis of a clear distinction in meaning, as in (4a vs. 4b, 4c vs. 4d), and when the different plural forms can be used interchangeably, and must be recognized as realizing the same lexeme, as in (5a-5b).

⁵Breton nouns inflect only for number.

⁶The French gloss given by Trépos (1980: 73) for *eskibien* is 'chevilles d'attelage'.

⁷Even if (1b) obviously derives from (1a) by means of a metaphorical extension.

(4) Modern Standard Arabic (Holes 2004, Kaye 2007)⁸

- a. SG *bayt* PL *buyu:t*
'tent', 'house'
- b. SG *bayt* PL *?abya:t*
'verse of poetry'
- c. SG *maktab* PL *maka:tib*
'office'
- d. SG *maktab* PL *maktaba:t*
'library', bookshop'

(5) Modern Standard Arabic (Kaye 2007)

- a. SG *ʕayn* PL *?aʕyun/ʕuyūn*
'eye'
- b. SG *sāriq* PL *sāriqūn, saraqa, surrāq*
'thief'

With respect to nouns such as those in (5), Kaye (2007: 234–235) observes that “[t]here are many nouns with two or more plural variants without any difference in meaning”, while on the nouns in (4a–4b) he states that “[i]t is best to regard [...] *bayt* as distinct lexemes” (Kaye 2007: 234).

Authors like Kaye rely on meaning distinction as the only criterion for distinguishing between lexemes, and (implicitly) accept the possibility that what they conceive of as single lexemes (like the ones in (5)) may have overabundant realizations in one or more cells, i.e., may map to more than one flexeme. Other authors, however, reject this possibility, and assume that a difference in inflectional realization (a difference in flexemes) must always correspond to a difference in lexemes. A champion of such a position is Paolo Acquaviva, who has articulated his point of view in his work on Italian double noun plurals (Acquaviva 2008).

Italian nouns have inherent gender (with two values: feminine and masculine) and inflect for number (with two values: singular and plural). About 20 Italian nouns are usually described as overabundant in the plural (e.g., in traditional reference grammars such as Battaglia & Pernicone 1954). These nouns have a singular form in *-o* which is masculine, a plural form in *-i* which is masculine, and a plural form in *-a* which is feminine. Some representative examples are given in (6):

(6) Italian (Acquaviva 2008, Thornton n.d.)

- a. SG *braccio* PL *braccia/bracci*
'arm'

⁸MS Arabic nouns inflect for number (singular, dual, plural), case (nominative, genitive, accusative, with a syncretism of genitive and accusative (sometimes called oblique) in non-singular forms), and definiteness (definite, indefinite). In systems in which nouns inflect for other features besides number, if multiple forms with the same number value exist they are predicted to exist in all cells; e.g., in Arabic, multiple plural forms are predicted to exist in all case and definiteness values.

- b. SG *corno* PL *corna/corni*
‘horn’
- c. SG *ginocchio* PL *ginocchia/ginocchi*
‘knee’
- d. SG *membro* PL *membra/membri*
‘limb’/‘member’

Acquaviva’s position is that plurals in *-a*, independently of whether they differ in meaning from the plurals in *-i* with which they share a root, are distinct lexemes, *pluralia tantum*, derivationally related to the lexemes in *-o/-i* with which they share a root:

plurals in *-a* [...] are lexical plurals: **distinct, inherently plural nouns**, related to the base noun by a word-formation process. (Acquaviva 2008: 123, emphasis mine)

***Braccia* ‘arms’ is not the plural of *braccio* ‘arm’**; it is an inherently plural lexeme, derived from the same root as *braccio/bracci* (Acquaviva 2008: 157, emphasis mine)

He brings forward several arguments for his position, which are reviewed in Thornton (n.d.: 430–438), where it is shown that one of them (agreement with conjoined singular NPs) is based on a misunderstanding of the workings of Italian agreement resolution rules, and can be dismissed as irrelevant. His other arguments will be illustrated here.

The first argument is purely metatheoretical. Acquaviva states it as follows:

The simple fact that a number of plurals in *-a* do not block their regular alternants in *-i* is enough to prove the point, **if we take seriously inflectional disjunctivity** (Acquaviva 2008: 145, emphasis mine).

This argument boils down to positing as a theoretical requirement the non-existence of overabundance, or the impossibility of a single lexeme to map to distinct flexemes. Such a choice eliminates the problem we are investigating by denying its existence, rather than by offering a solution. However, if we assume, as done in the canonical approach to morphological typology (Corbett 2005, 2006, 2007), that inflectional disjunctivity and lack of overabundance are only canonical properties of lexemes, rather than inviolable theoretical requirements, the problem reappears and requires to be investigated.

Another argument put forward by Acquaviva to establish that plurals in *-a* are distinct lexemes from their co-radicals in *-o/-i* is consonant with Fradin & Kerleroux’s (2003) view of lexemes: Acquaviva observes that some plurals in *-a* appear to be the bases of word-formation processes. An example would be *cornificare* ‘to make a cuckold of’, which Acquaviva analyzes as derived from *corna* ‘horns’ (6b); *cornificare* is synonymous with the idiom *fare/mettere le corna* ‘to make a cuckold of, lit. to make /put horns.F.PL’, which is never realized by **fare/mettere i corni*, with ‘horns.M.PL’. On this basis, one can presume that *corna*, and not *corni*, is the base of *cornificare*. However, the idiom *fare/mettere un corno* ‘to make a cuckold of, lit. to make/put a horn.M.SG’ is also attested, so one cannot exclude that the base of *cornificare* is a non-defective lexeme *corno/corna*, rather than

a *plurale tantum* defective noun *corna*. In any case, this argument boils down to recognizing different lexemes when there is a difference in semantics and in the possibility of appearing in certain constructions, as proposed also by Fradin & Kerleroux (2003, 2009), Fradin (2003). This is orthogonal to the question whether a lexeme, defined on the basis of its semantics and distribution in constructions, can be overabundant in one or more cells. If we show that two plural forms appear in the same set of environments and constructions, they must be recognized as belonging to the same lexeme (unless, like Acquaviva, one wants to posit a difference of inflectional realization as sufficient for recognizing distinct lexemes, regardless of the equal semantics and distribution of the forms). Thornton (2010-2011) has shown, by means of corpus-based evidence, that in some cases two plurals in *-i* and *-a* are used interchangeably in the same context, and cannot therefore be considered as instances of distinct lexemes in Fradin & Kerleroux (2003)'s sense. This is the case for *ginocchi* / *ginocchia* 'knees' (6c), both of which appear interchangeably (as well as the singular form *ginocchio*) in a number of syntactic environments (Thornton n.d.: 465). In the case of *membra* and *membri* (6d), instead, there is evidence to posit two distinct lexemes, MEMBRO¹ 'limb (body part)', which is [-human], and MEMBRO² 'member (of a committee, organization, etc.)', which is [+human], and is obviously derived from MEMBRO¹ by means of a metaphoric extension. MEMBRO² is not overabundant: its plural is always *membri*, and it is the base of a derived feminine MEMBRA 'female member (of a committee, organization, etc.)', PL *membre* (Thornton 2014). MEMBRO¹ isn't overabundant either: its plural is *membra* 'limbs'; however, contrary to Acquaviva's analysis, it is not defective: the singular *membro* in the sense of 'limb, body part' is attested (cf. Thornton n.d.: 463, fn. 38). These examples show that each case in which we observe, in Italian, a feminine plural in *-a* and a masculine one in *-i* based on the same root, must be analyzed in its own right: the parallelism in the flexemes does not guarantee a parallelism in the lexemes. *Membri* and *membra* belong to different lexemes (defined according to Fradin & Kerleroux's (2003) and Fradin's (2003) semantic criteria), while *ginocchi* and *ginocchia* belong to the same lexeme – if we admit the possibility of overabundance, i.e. of a single lexeme mapping to more than one flexeme. The case of *bracci* and *braccia* is particularly complex: these very frequent forms, if submitted to Fradin & Kerleroux's (2003) and Fradin's (2003) criteria for the recognition of distinct lexemes, map to several semantically distinct lexemes, some of which are overabundant in the plural (e.g., 'arm (body part)'), while others select only one plural form (e.g., 'ell (measure of length)' selects *braccia*). Again, the mapping between lexemes and flexemes is not 1:1, as shown in Figure 1.

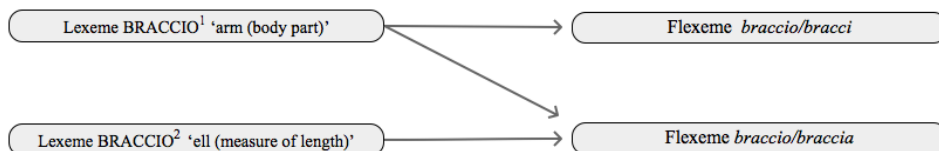


Figure 1: Mapping between two lexemes and two flexemes in Italian

4.2 Case study 2: Past participles

Another area in which mapping between semantically defined lexemes and flexemes is not always 1:1, and in which differences in flexemes do not invariably coincide with differences in lexemes, is verbal inflection.

In some cases, two semantically and constructionally distinct lexemes have quite different realized paradigms, even if the citation forms coincide. A case in point is that of Italian *SUCCEDERE*¹ ‘happen’ and *SUCCEDERE*² ‘succeed’. *SUCCEDERE*¹ ‘happen’ is an impersonal verb, which is used only in 3rd person forms; its PST.PTCP is *successo*. *SUCCEDERE*² ‘succeed’ is a bi-argumental verb; its second argument is introduced by the preposition *a* ‘to’; it has a full set of realized forms, and its PST.PTCP is overabundant, according to various authoritative sources (Zingarelli 2016, Serianni 1988): it can be either *succeduto* or *successo*. The forms are shown in (7):

- (7) Italian (Zingarelli 2016, Serianni 1988, personal knowledge)
- | lexeme | PST.PTCP |
|---|---------------------------|
| <i>SUCCEDERE</i> ¹ ‘happen’ | <i>successo</i> |
| <i>SUCCEDERE</i> ² ‘succeed’ | <i>successo/succeduto</i> |

From (7) it would appear that *SUCCEDERE*¹ maps to a single flexeme, while *SUCCEDERE*² maps to two. However, for *SUCCEDERE*² the form *succeduto* is prescribed over *successo* by normatively oriented sources like Serianni (1988: § 316), and the most recent example of *successo* as a form of *SUCCEDERE*² cited by Serianni (1988) is from a novel published in 1960. Investigation of contemporary usage in corpora is difficult for practical reasons: *successo* has 87763 tokens in the corpus *la Repubblica 1985-2000* (380M tokens; I will consider data from this corpus as representative of contemporary Italian usage of *successo* and *succeduto*), making it impractical to examine each token to assign it to either *SUCCEDERE*¹ or *SUCCEDERE*². Besides, *successo* is a homonym of the SG form of the noun *SUCCESSO* ‘success’. However, manual examination of the first 200 random tokens of the string *successo a*, which corresponds to both ‘happened to’ and ‘succeeded to’, suggests that in this context *successo* always realizes *SUCCEDERE*¹ ‘happen’, while, as expected, all the 374 tokens of *succeduto* in the corpus *la Repubblica 1985-2000* realize *SUCCEDERE*² ‘succeed’. So, as far as the PST.PTCP is concerned, it appears that in contemporary Italian the two lexemes *SUCCEDERE*¹ ‘happen’ and *SUCCEDERE*² ‘succeed’ map to different flexemes.⁹

We can compare this situation with that of the verb *PERDERE* ‘lose’, which is genuinely overabundant in its PST.PTCP, as shown in (8):

⁹Things are more complicated with the simple past, which is (exemplifying with 3SG forms) *successe* for *SUCCEDERE*¹ and overabundant for *SUCCEDERE*² (*successe/ succedette*; a third form, *succedé*, is theoretically possible as ‘succeed.PST.3SG’, but it is not attested in the corpus *la Repubblica 1985-2000*). *Successe* has 1263 tokens and *succedette* 43 tokens in this corpus; all tokens of *succedette* realize *SUCCEDERE*² ‘succeed’; manual examination of the 14 tokens of the string *successe a* ‘happened to/succeeded to’ reveals that in most cases it realizes *SUCCEDERE*¹ ‘happen’, but in 2 cases *successe* realizes *SUCCEDERE*² ‘succeed’, confirming that this verb is overabundant in its simple past. However, the simple past does not belong to the native grammar of many speakers of Italian, for whom it is a learned form; so it is unwise to draw strong conclusions from these data. Overabundance in the simple past in Italian shall be left for further research.

- (8) Italian (personal knowledge)
lexeme PST.PTCP
perdere 'lose' *perso/perduto*

Speakers appear unaware of conditions regulating the selection of either one of the two forms, to the point that many speakers asked the *Accademia della Crusca's* linguistic consulting service for advice on when to use each form (Thornton 2016). Speakers seem convinced that rules that govern a complementary distribution of the two forms should exist, but indeed the distribution of the two PST.PTCP forms is not complementary: they can be used interchangeably in many contexts, including idioms, as shown in (9a-b) and already shown by Thornton (2011: 369); the only case in which only one form is used is in titles of works of art (9c). Representative data, with frequencies from the corpus *la Repubblica 1985-2000* when relevant, are presented in (9).

- (9) Italian (Thornton 2011: 369, Thornton 2016, personal knowledge)
- a. *occasione perduta* 291 / *occasione persa* 83
'a chance lost'
 - b. *perso la guerra* 109 / *perduto la guerra* 32
'lost the war'
 - c. *I predatori dell'arca perduta*/**persa*
'Raiders of the lost ark'
 - d. *Alla ricerca del tempo perduto*/**perso*
À la recherche du temps perdu by Proust, literally 'In search of lost time';
English translation's title 'Remembrance of things past'
 - e. *Paradiso perduto*/**perso*
'Paradise lost'

This case study shows again a case in which similar differences in flexemes do not map in a parallel way to differences in lexemes: while *SUCCEDERE*¹ 'happen' and *SUCCEDERE*² 'succeed' map to distinct flexemes, in which the PST.PTCP forms are *successo* and *succeduto* respectively, *PERDERE* 'lose' maps to two flexemes, distinct from each other in a way parallel to the flexemes *SUCCEDERE*¹ and *SUCCEDERE*², and its PST.PTCP can be realized by both *perso* and *perduto*.

4.3 Systematic overabundance and overabundance in all cells

The two case studies illustrated above have shown examples in which there is an overabundant cell in the form paradigm and the realized paradigm of certain lexemes (such as Italian *BRACCIO*¹ 'arm' and *PERDERE* 'lose'). Technically, this should be enough to recognize that such lexemes map to distinct flexemes. However, if one wished to take into

account quantitative considerations, one might want to deal with these cases by recognizing a minor “exception”, and still posit a single flexeme with a single exceptional, overabundant cell.

However, overabundance is not always confined to a single cell. In this section I will illustrate cases of “systematic overabundance” (Bonami & Stump 2016: 469), in which entire slabs or subparadigms are involved,¹⁰ and cases of overabundance in all cells. These cases definitely deserve consideration in the context of exploring the possible deviations from a 1:1 mapping between lexemes and flexemes.

A particularly clear example of systematic overabundance is found in Spanish, where all verbs have two complete sets of forms, built by means of different endings, in the Imperfect Subjunctive, as shown in Table 1 for the verb *haber* ‘have’.

Table 1: Imperfect Subjunctive of Spanish *haber* ‘have’

	<i>-ra</i> set	<i>-se</i> set
1SG/3SG	<i>hubiera</i>	<i>hubiese</i>
2SG	<i>hubieras</i>	<i>hubieses</i>
1PL	<i>hubiéramos</i>	<i>hubiésemos</i>
2PL	<i>hubiérais</i>	<i>hubiéseis</i>
3PL	<i>hubieran</i>	<i>hubiesen</i>

Despite a suggestion by Bolinger (1956) that there is some subtle semantic difference between the two sets of forms, contemporary descriptions agree that “these two sets of forms are interchangeable” (Butt & Benjamin (2000: 167); see also Rojo & Veiga (1999: 2910): “las formas en *-ra* y *-se* son hoy por hoy perfectamente equivalentes”). Spanish verbal lexemes, then, appear to systematically map to two flexemes, which are distinct in the Imperfect Subjunctive forms – unless one wants to build overabundance within the definition of Spanish verbal flexemes, exactly because of its systematicity.

In other cases, however, we encounter overabundance in all cells of a given lexeme, but this is not systematic across all the lexemes within that part of speech in the language; therefore, the possibility of building overabundance in the definition of the flexemes to which these lexemes map is not viable, and we must recognize a 1:2 mapping between lexemes and flexemes.

A case in point is that of the Italian noun *ORECCHIO* ‘ear’. This noun can be described as overabundant in all its cells: it has two SG forms and two PL forms, as shown in (10):

¹⁰The notion of slab has been introduced by Carstairs (1987: 81), who defines it as “a subset of the macroinflections within one paradigm consisting of all the macroinflections which are associated with some specified morphosyntactic property”. His examples from Latin noun paradigms are the singular slab (all singular case-forms) or the genitive slab (GEN.SG and GEN.PL). The notion of sub-paradigm is used in a variety of senses, most commonly by scholars with a background in Slavonic languages. It aims at capturing subsets of cells in a paradigm which share more than just one feature value, such as verb tenses (the Present Indicative, the Present Subjunctive, etc.).

- (10) Italian (personal knowledge)
 lexeme ORECCHIO
 SG forms *orecchio* (M) *orecchia* (F)
 PL forms *orecchi* (M) *orecchie* (F)

Of course, one could posit two distinct lexemes, ORECCHIO(M) and ORECCHIA(F), on the basis of the difference in gender, which is canonically an inherent fixed feature value in nouns. However, we already know from the cases discussed in Section 4.1 that Italian has nouns which change their gender value from the singular to the plural. Besides, according to Fradin’s (2003) and Fradin & Kerleroux’s (2003) definition of lexeme, which recognizes a single lexeme on the basis of identity of meaning and constructional distribution, the different forms in (10) appear to belong to the same lexeme, since they can be used interchangeably in the same contexts, even in idioms (11a-11b), as shown by the examples in (11):

- (11) Italian (personal knowledge; frequency data from the corpus *la Repubblica 1985-2000*)
- a. *fare orecchi da mercante* 18 / *orecchie da mercante* 139
 ‘to turn a deaf ear’ lit., to do merchant’s ears
 - b. *dare una tirata d’orecchi* 122 / *tirata d’orecchie* 92
 ‘to give a dressing-down’ lit., to give a tug of ears
 - c. *occhi e orecchi* 19 / *occhi e orecchie* 68
 ‘eyes and ears’
 - d. *da un’orecchia all’altra* 2 / *da un’orecchio all’altro* 13
 ‘from one ear to the other’

So Italian ORECCHIO can be analyzed as a single lexeme mapping to two flexemes, as shown in Figure 2.

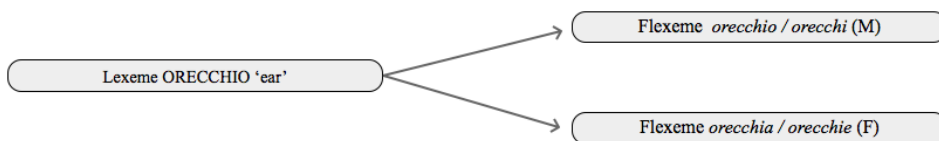


Figure 2: Mapping between one lexeme and two flexemes in Italian.

The flexemes are distinct; they instantiate nouns of different inflectional classes, while most Italian noun lexemes map to only one flexeme, belonging consistently to only one gender and one inflectional class, as shown by the examples in Table 2.

Lexemes such as BRACCIO¹, GINOCCHIO and ORECCHIO are non-canonical, in that they map to more than one flexeme, as seen above.

Table 2: Italian (personal knowledge).

lexeme	flexeme		gloss
	SG	PL	
OCCHIO (M)	<i>occhio</i>	<i>occhi</i>	‘eye’
BOCCA(F)	<i>bocca</i>	<i>bocche</i>	‘mouth’
MANO(F)	<i>mano</i>	<i>mani</i>	‘hand’

The last case of non-canonical mapping between lexemes and flexemes that I will examine is that of certain Italian verbs, that are described as able to inflect according to two different conjugations; these are called “verbi sovrabbondanti” by Serianni (1988).

Grammars usually address together two kinds of such verbs: those in which the difference in conjugation does not bring along a difference in meaning (12a), and those in which the difference in inflectional class goes hand in hand with a difference in meaning (12b).

(12) Italian (Serianni 1988, personal knowledge)

- a. i. *adempiere/adempire*
‘fulfil’
- ii. *compiere/compire*
‘complete’
- iii. *empiere/empire*
‘fill’
- iv. *riempiere/riempire*
‘fill’
- b. i. *abbonare/abbonire*
‘subscribe’/‘appease’
- ii. *arrossare/arrossire*
‘make red’, ‘dye red’/‘redden’, ‘flush’
- iii. *fallare/fallire*
‘make a mistake’/‘fail’
- iv. *imboscare/imboschire*
‘hide [in a wood]’/‘afforest’
- v. *impazzare/impazzire*
‘be in full swing’/‘go crazy’
- vi. *sfiurare/sfiurare*
‘brush’, ‘graze’/‘wither’, ‘wilt’

Serianni (1988), from which the examples in (12) are taken, considers the cases in (12a) and (12b) as two groups of overabundant verbs, while Dardano & Trifone (1985) consider only cases (12a) as overabundant verbs, and propose that cases in (12b) are best analyzed as distinct lexemes; I concur with Dardano & Trifone, because of a clear difference in meaning between the two verbs in each pair in (12b); these verbs are different lexemes according to Fradin's (2003) and Fradin & Kerleroux's (2003) criteria, and will not be further discussed here.

Verbs in (12a) are claimed to have forms belonging to the two inflectional classes traditionally called 2nd conjugation (infinitive ending in *-ere*) and 3rd conjugation (infinitive ending in *-ire*); besides, the 3rd conjugation forms belong to the subclass of 3rd conjugation verbs which does not exhibit the element *-isc-* in the appropriate morphomic partition (so PRS.IND.1SG is *empio*, not **empisco*, etc.). The 2nd conjugation and the *-isc-*less subclass of the 3rd conjugation have non-distinct inflection in several cells, listed in (13a), while they have distinct forms in other cells, listed in (13b), with examples from *riempiere* and *riempire*.¹¹

(13) Italian (personal knowledge)

- a. Cells with non-distinct realization for the verbs in (12a)
Present Indicative: all person/number forms, except 2PL
Present Subjunctive: all person/number forms
Imperative 2SG
Gerund
(Present Participle)¹²
- b. Cells with distinct realization for the verbs in (12a)
Present Indicative 2PL = Imperative 2PL (e.g., *riempiete* vs. *riempite*)
Imperfective Past Indicative (*Imperfetto*): all person/number forms (e.g., 1SG *riempievo* vs. *riempivo*, etc.)
Simple Perfective Past Indicative (*Passato Remoto*): all person/number forms (e.g., 1SG *riempieetti* or *riempieii* vs. *riempii*, etc.)
Future: all person/number forms (e.g., 1SG *riempierò* vs. *riempirò*, etc.)
Imperfect Subjunctive: all person/number forms (e.g., 1SG *riempiessi* vs. *riempissi*, etc.)

¹¹In (13) I consider only synthetic forms; periphrastic forms are formed by an inflected auxiliary followed by a Past Participle, so their distinctness is a function of the distinctness of the Past Participle form (therefore, they are always distinct for these two conjugations).

¹²A so-called Present Participle ending in *-nte* is normally listed as part of a verb's paradigm in Italian descriptive grammars, but it is extremely doubtful that such a cell should be recognized as a genuine part of verbal paradigms in Italian. Haspelmath (1996) contrasts these so-called present participles of Italian with those of other languages in terms of their syntactic properties (government of subject and non-subject arguments) and concludes that in Italian "active participles do not exist" (Haspelmath 1996: 61). Luraghi (1999) is less drastic, but shows that *-nte* forms have never been part of the spoken register in the history of the language, and that a verbal usage of *-nte* forms is only attested in some technical or bureaucratic registers, while adjectives and nouns in *-nte*, often unrelated to any verbal base, are common.

Present Conditional: all person/number forms (e.g., 1SG *riempirei* vs. *riempirei*, etc.)

Past Participle (e.g., *riempito* vs. *riempito*)

Infinitive (e.g., *riempire* vs. *riempire*)

The verbs in (12a) are technically cases of single lexemes mapping to two distinct flexemes, but these flexemes are syncretic in all the cells listed in (13a).

As I am always wary of believing statements by grammars on the distribution of cell mates, I have checked the distribution in the corpus *la Repubblica 1985-2000* of the forms of the verbs in (12a) that are distinct in the two conjugations. Table 3 illustrates the results (figures for forms of the same Tense/Mood have been added together).

The data in Table 3 show the following picture: EMPIERE/EMPIRE ‘fill’ are almost extinct verbs in both conjugations, totaling only 13 forms overall; their meaning is normally expressed, in contemporary Italian, by RIEMPIERE; RIEMPIERE ‘fill’ is little used – there are a few tokens of the Infinitive and of the Imperfective Past Indicative (*Imperfetto*) in usage, but the ratio between forms of RIEMPIERE and forms of RIEMPIERE in the cells for which the two conjugations have distinct forms is so unbalanced (504:1) that the two verbs represent at best an extremely weak and non-canonical case of overabundance (or mapping from one lexeme to two flexemes) according to Thornton’s (2012: 188–189) criteria for measuring the strength of overabundance on the basis of frequency ratios between two cell mates. ADEMPIERE and ADEMPIRE ‘fulfil’ have a less unbalanced frequency ratio (15.2:1) overall, but it must be observed that 99.5% of the forms of ADEMPIERE are realizations of the Infinitive and the Past Participle, while 93.3% of the forms of ADEMPIRE are realizations of tenses different from the Infinitive and the Past Participle. Indeed, all the Past Participle forms are 2nd conjugation forms (i.e., they are forms of ADEMPIERE, not possible forms of ADEMPIRE), so there is no overabundance in this cell; the only tenses in which the two verbs display some overabundance are the Future (with a ratio of 5.4:1 in favour of ADEMPIRE) and, very marginally, the Infinitive (with a very unbalanced ratio of 154:1 in favour of ADEMPIERE). The same picture, even more dramatically, is presented by COMPIERE/COMPIRE ‘complete’. Assessment of overabundance in this case is made difficult by the fact that some Past Participle forms of COMPIERE are homographous with other forms in the paradigm, and/or with forms of the noun COMPITO ‘task, homework’, and/or of the adjective COMPITO ‘courteous, polite’, and/or of the verb COMPITARE ‘spell out’ (e.g., *compito* represents ‘complete.PST.PTCP.M.SG’, ‘task(M).SG’, ‘courteous.M.SG’ and ‘spell_out.PRS.IND.1SG’; the noun for ‘task’ and the 1SG form of ‘spell out’ have antepenultimate stress, while the other forms have penultimate stress, but stresses on these syllables are not marked in the standard orthography of Italian, so all the forms are homographs even if they are not all homophonous); these homographies have been manually disambiguated for the forms ending in *-a* and *-e* (*compita* ‘complete.PST.PTCP.F.SG’, ‘courteous.F.SG’, ‘spell_out.PRS.IND.3SG’ and *compite* ‘complete.PST.PTCP.F.PL’, ‘complete.PRS.IND.2PL’, ‘courteous.F.PL’), which have low frequency, thus making manual disambiguation practical; the lack of manual disambiguation for the high frequency forms in *-o* and *-i* explains why the exact frequency of these forms is

Table 3: Frequency in *la Repubblica 1985-2000* corpus of forms of the verbs in (12a) in cells that have distinct realizations for the two conjugations.

	EMPIERE	EMPIRE	RIEMPIERE	RIEMPIRE	ADEMPIERE	ADEMPIRE	COMPIERE	COMPIRE
2PL.PRS.IND = 2PL.IMP	--	1	--	368	--	2	--	2
<i>Imperfetto</i>	--	2	2	611	--	10	7	441
<i>Passato Remoto</i>	1	5	--	282	--	7	--	423
Future	1	--	--	294	5	27	23	1371
Imperfect Subjunctive	--	--	--	52	--	7	7	111
Present Conditional	--	--	--	52	--	3	3	76
Past Participle	--	--	--	2307 -o	233 -o	--	15907 -o	? -o
				537 -a	10 -a		4700 -a	3 -a
				450 -i	24 -i		5926 -i	? -i
				368 -e	24 -e		2397 -e	1 -e
Infinitive	--	3	15	3239	615	4	9267	8
Total	2	11	17	8560	911	60	38237	> 2400

not given in Table 3, and a question mark has been inserted instead.¹³ The actual forms realizing ‘complete.PST.PTCP.F.SG’, and ‘complete.PST.PTCP.F.PL’ turned out to be a minority (3 over 48 (6%) for the F.SG form, 1 over 9 (11%) for the F.PL form). Therefore, it may be concluded with some safety that in the Past Participle cell the verb *COMPIERE* is favoured and *COMPIRE* is quite underrepresented. These two verbs show the same kind of “division of labour” already observed for *ADEMPIERE* and *ADEMPIRE*: *COMPIERE* specializes for the Infinitive and the Past Participle, and *COMPIRE* for all other tenses (among the ones that have distinct realizations for the two conjugations); however, in most tenses a few forms of *COMPIERE* are also attested, so *COMPIERE/COMPIRE* represent the best example of overabundance in all cells encountered so far among the Italian verbs commonly dubbed “sovrabbondanti” (although the frequency ratios render this case of overabundance not very canonical). It seems that *ADEMPIERE/ADEMPIRE* and *COMPIERE/COMPIRE* are on their way from overabundance to heteroclysis: at some point in the future, we might observe a lexeme with finite synthetic forms belonging to the 3rd conjugation and Infinitive and Past Participle (which carries with it all the periphrastic forms) belonging to the 2nd conjugation. *RIEMPIERE/RIEMPIRE*, instead, is just reducing overabundance in favour of the 3rd conjugation forms, and is quite advanced in this process.

If the process leading to heteroclysis is completed, we will have a single lexeme mapping to a single heteroclitic flexeme. At the moment, however, we have a number of Italian verbal lexemes that map to two flexemes, at least in parts of their paradigm.¹⁴

5 Conclusions

The data illustrated in this paper show that the distinction between lexemes and flexemes first proposed by Fradin & Kerleroux (2003) and Fradin (2003), as well as their definition of lexeme based on semantic and constructional coherence, is useful even beyond the area of lexeme formation, for which it was originally proposed. A separation between lexemes and flexemes, like the separation between content paradigms, form paradigms and realized paradigms adopted in paradigm-linkage theory, is a useful tool in models of morphological analysis that recognize a level of autonomous morphology.

References

- Acquaviva, Paolo. 2008. *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge: MIT Press.

¹³The raw frequency in *la Repubblica 1985-2000* of the form *compito* is 26450, that of *compiti* 9180.

¹⁴A reviewer observes that one could introduce overabundance in the very make-up of flexemes, rather than using it as the grounds for positing two (or more) distinct flexemes whenever a lexeme has two (or more) distinct realizations for a single cell, as done in this paper. This would involve allowing flexeme cells to comprise sets of forms rather than a single form. The relative merits of the two alternative approaches could be fully compared only in a formalized model, whose development exceeds the scope of this paper.

- Aronoff, Mark. 2000. Morphology between lexicon and grammar. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphology. An international handbook on inflection and word-formation*, vol. 1, 344–349. Berlin: Walter de Gruyter.
- Battaglia, Salvatore & Vincenzo Pernicone. 1954. *La grammatica italiana. Seconda edizione migliorata*. Turin: Loescher Chiantore.
- Bolinger, Dwight. 1956. Subjunctive *-ra* and *-se*: “Free Variation”? *Hispania* 39(3). 345–349.
- Bonami, Olivier & Gregory T. Stump. 2016. Paradigm Function Morphology. In Andrew Hippisley & Gregory T. Stump (eds.), *The Cambridge Handbook of Morphology*, 449–481. Cambridge: Cambridge University Press.
- Bonami, Olivier & Delphine Tribout. 2012. Underspecification and the semantics of lexeme formation. Vienna. <http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/biblio//diapositive-v5.pdf>, accessed 2016-11-06. Paper presented at the 15th International Morphology Meeting.
- Butt, John & Carmen Benjamin. 2000. *A new reference grammar of modern Spanish*. 3rd edition. London: Arnold.
- Carstairs, Andrew. 1987. *Allomorphy in inflexion*. London: Croom Helm.
- Corbett, Greville G. 2005. The canonical approach in typology. In Zygmunt Frajzyngier, Adam Hodges & David S. Rood (eds.), *Linguistic diversity and language theories*, 25–49. Amsterdam: John Benjamins.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2007. Canonical typology, suppletion, and possible words. *Language* 83(1). 8–42.
- Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Max Niemeyer Verlag.
- Dardano, Maurizio & Pietro Trifone. 1985. *Grammatica italiana con nozioni di linguistica*. Bologna: Zanichelli.
- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.
- Fradin, Bernard. Forthcoming. Competition in derivation: What can we learn from duplicates? In Francesco Gardani, Hans Christian Luschützky, Franz Rainer & Wolfgang U. Dressler (eds.), *Competition in morphology*. Dordrecht: Springer.
- Fradin, Bernard & Françoise Kerleroux. 2003. Troubles with lexemes. In Geert Booij, Janet DeCesaris, Angela Ralli & Sergio Scalise (eds.), *Selected papers from the third Mediterranean Morphology Meeting*, 177–196. Barcelona: IULA – Universitat Pompeu Fabra.
- Fradin, Bernard & Françoise Kerleroux. 2009. L’identité lexémique. In Bernard Fradin, Françoise Kerleroux & Marc Plénat (eds.), *Aperçus de morphologie du français*, 83–102. Saint-Denis: Presses Universitaires de Vincennes.
- Haspelmath, Martin. 1996. Word-class-changing inflection and morphological theory. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1995*, 43–66. Dordrecht: Kluwer Academic Publishers.

- Holes, Clive. 2004. *Modern Arabic. Structures, functions and varieties*. Washington, D. C.: Georgetown University Press.
- Kaye, Alan S. 2007. Arabic morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, 211–247. Winona Lake: Eisenbrauns.
- la Repubblica 1985–2000 *corpus*. N.d. <http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica>. (Accessed 6 November 2016).
- Luraghi, Silvia. 1999. Il suffisso *-ante/-ente* in italiano: fra flessione e derivazione. In Paola Benincà, Alberto M. Mioni & Laura Vanelli (eds.), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso della Società di linguistica italiana*, 539–550. Rome: Bulzoni.
- Rojo, Guillermo & Alexandre Veiga. 1999. El tiempo verbal. Los tiempos simples. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. 2, 2867–2934. Madrid: Espasa Calpe.
- Serianni, Luca. 1988. *Grammatica italiana*. Turin: Utet. In cooperation with Alberto Castelvechi.
- Stump, Gregory T. 2016. *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
- Thornton, Anna M. N.d. La non canonicità del tipo it. *braccio // braccia / bracci*: Sovrabbondanza, difettività o iperdifferenziazione? *Studi di grammatica italiana* 29–30. 419–477. Published in 2013.
- Thornton, Anna M. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology. In Martin Maiden, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds.), *Morphological autonomy: Perspectives from Romance inflectional morphology*, 358–381. Oxford: Oxford University Press.
- Thornton, Anna M. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5(2). 183–207.
- Thornton, Anna M. 2014. Risposta n. 7 [le membre del comitato]. *La Crusca per voi* 49(II). 14–15.
- Thornton, Anna M. 2016. *Perso o perduto? Visto o veduto?* <http://www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica/domande-risposte/perso-perduto-visto-veduto>. (Accessed 31 October 2016).
- Trépos, Pierre. 1980. *Grammaire bretonne*. Rennes: Ouest France.
- Ullmann, Stephen. 1957. *The principles of semantics*. Oxford: Basil Blackwell & Mott.
- Zingarelli. 2016. *Lo Zingarelli 2016. Vocabolario della lingua italiana di Nicola Zingarelli*. Bologna: Zanichelli. Ed. by Mario Cannella and Beata Lazzarini.

