

# Morphome interactions

Borja Herce<sup>1,2</sup> 

Received: 19 June 2018 / Accepted: 26 November 2018

© The Author(s) 2018

**Abstract** How morphological elements interact with one another is a major concern for both morphological theory and typological research. Morphemes, for example, are usually said to block each other when they are in subset-superset relations to one another. Little is known, by contrast, of how morphomes (i.e. forms with unnatural morphosyntactic distributions) interact with one another. This paper provides an initial typology of morphome interactions, based on whether their forms overlap in the paradigm or not, or whether the distribution of one constitutes a subset of the other or not. In addition, special attention will be paid to the outcome or the resolution of morphome orthogonalities. Focusing on the interaction between the L- and N-morphemes of Romance, the analogical changes related to their orthogonality are surveyed and a constraint is proposed that limits the possibilities of morphome interaction.

**Keywords** Morphome · Romance · Typology · N-morphome · L-morphome · Paradigm · Analogy

## 1 Introduction

Morphomes (Aronoff 1994), or metamorphomes in Round's (2013) terminology, are systematic patterns of morphological identity whereby forms are not associated with any natural morphosyntactic or semantic class. They constitute autonomously morphological patterns that often originate as the accidental product of sound change but which are subsequently learnt, replicated, reinforced and preserved in diachrony in a way which suggests that they represent generalizations or categories that are cognitively real for language users.

---

✉ B. Herce  
[borjaherce@gmail.com](mailto:borjaherce@gmail.com)

<sup>1</sup> University of the Basque Country (UPV/EHU), Vitoria-Gasteiz, Spain

<sup>2</sup> University of Surrey (SMG), Guildford, UK

**Table 1** Koyi Rai *jɔmts-/jɔmd-* ‘hit’ (Baerman and Corbett 2012, from Lahaussois 2009:12)

	SG	DU	PL
1EX	jɔmd-uŋa	jɔmts-asu	jɔmts-ako
1INC		jɔmts-asi	jɔmts-aki
2	jɔmd-una	jɔmts-asina	jɔmts-ani
3	jɔmd-u	jɔmd-usi	jɔmd-uni

**Table 2** Menngwa Dla *numb-/nung-* ‘stand’ (de Sousa 2006:541)

	SG	DU	PL
1	numb-aha-hwa	numb-ehya-hwa	numb-efa-hwa
2M	numb-afa-hwa	numb-afa-hwa	nung-uma-hwa
2F	numb-afa-hwa	numb-efya-hwa	numb-ei-hwa
3M	nung-u-hwa	numb-afa-hwa	nung-uma-hwa
3F	nung-wa-hwa	numb-efya-hwa	numb-ei-hwa

There is by now a substantial literature on the phenomenon (e.g. Maiden 1992; O’Neill 2011; Maiden et al. 2011; Luís and Bermúdez-Otero 2016 etc.). For reasons of coverage and tradition, most of the literature on the morpheme has been so far been concerned with data from Romance. However, the phenomenon is by no means exclusive to any one language family, of course. Consider, for example, the following less familiar cases as similarly representative of what (meta)morphemes are as shown in Tables 1 and 2.

In Koyi Rai (Kiranti, Sino-Tibetan) we can see how singular and 3rd person share a stem different from the one found in the rest of the paradigm. In Menngwa Dla (Senagi, Papua), the same stem is used for 3SG and for 2/3PL.M. These classes do not make sense from a morphosyntactic perspective. Despite their unnaturalness, both patterns are highly systematic, as they are replicated in various other lexemes and with very different formal correlates (e.g. *war-/ward-* ‘throw’, *hoʔ-/huʔ-* ‘bring’, *pja-/pa-* ‘eat’ etc. in Koyi Rai, and *eh-/s-* ‘talk’, *ah-/s-* ‘think’/‘call’, *ap-/e-* ‘sleep’ etc. in Menngwa Dla.).

These are the kind of morphological structures that, probably because of the predictive relations they afford (Blevins 2016:105), serve as useful generalizations to language users, who, in turn, perpetuate the inherited alternations and even extend the pattern to new contexts. A closer look at the Koyi Rai and Menngwa Dla patterns will reveal that the presence of the shaded stem alternants appears to correlate with those phonological environments before a high back vowel /u/ or a glide /w/. Even if there is, synchronically, no phonological rule in the languages that would account for the observed alternations, it is clear that the different phonological contexts are the most likely diachronic source for the patterns.<sup>1</sup> Because there is nothing extraordinary about how morphomic patterns emerge, there is nothing that prevents them from

<sup>1</sup>Note, however, that, generally, trying to account for stem alternations with reference to the shapes of affixes is little more than pushing the burden of explanation somewhere else. If we did that here, for example, the distribution of ‘a’ and ‘u’-initial suffixes in Koyi Rai would still remain arbitrary.

**Table 3** Fictional Koyi Rai ‘hit’ (modified from Table 1)

	SG	DU	PL
1EX	jɔbd-uŋa	jɔmts-asu	jɔmts-ako
1INC	jɔbd-uŋa	jɔmts-asi	jɔmts-aki
2	jɔbd-una	ɔbts-asina	jɔbts-ani
3	jɔmd-u	jɔmd-usi	jɔbd-uni

arising various times in the history of a language. Imagine, for example, that in Koyi Rai, some sound change denasalized a nasal before another one in the same word (i.e. jɔmd-uŋa > jɔbd-uŋa) (see Table 3).

The result would be a much more complex exponence, especially when the patterns overlap in the paradigm, as in Table 3. The number of stem alternants is multiplied (from two to four) and the conditioning environments for stem alternation are obscured. This is likely to result in substantial empirical differences in the behaviour of competing, multiple or cross-classifying morphomic exponents with respect to “simple” morphomic exponences like those in Tables 1 and 2. However, and despite initial approaches in works like Maiden (2012), Maiden (2018:Chap. 10) and O’Neill (2018) these cases have not been systematically analyzed and typologized. This is the purpose of the present paper. Because most of the data that will be presented throughout this paper will be from Romance, Sect. 2 will briefly introduce the reader to Romance morphomes. Section 3, in turn, will deal with the possible logical configurations of two morphomes in a paradigm. Section 4 will present the overlap of two such morphomes (the N- and L-morphomes) in Romance and the ways in which the conflict leads to new morphomes or to analogical processes that target some cell combinations but not others (a constraint will be proposed in this respect). To conclude, Sect. 5 will summarize the findings and will propose future avenues for research.

## 2 Introduction to the morphomes of Romance

Exactly the multiple morphome emergence that we fictionally envisaged in the previous section has actually taken place in the Romance family. The present section is a most succinct introduction to the historical origin of these morphological structures aimed at making subsequent discussion accessible to those without an insider knowledge of the topic. For more detailed and complete description the reader is referred to Maiden (2018) and the rest of the literature (e.g. Maiden 2004; O’Neill 2011) that will be cited here.

### 2.1 The N-morphome

Stress placement was predictable in Latin.<sup>2</sup> However, in the verbal system (i.e. when applied to person-number and TAM-suffixed forms) it meant that some verb forms

<sup>2</sup>The second-to-last syllable was stressed if that syllable was long. Otherwise, the third-to-last syllable was stressed.

**Table 4** Two illustrative examples of the Romance N-morphome

	Spanish <i>poder</i> ‘be able to’				Italian <i>andare</i> ‘go’			
	Indicative		Subjunctive		Indicative		Subjunctive	
	SG	PL	SG	PL	SG	PL	SG	PL
1	puedo	podemos	pueda	podamos	vado	andiamo	vada	andiamo
2	puedes	podéis	puedas	podáis	vai	andate	vada	andiate
3	puede	pueden	pueda	puedan	va	vanno	vada	vadano

(the shorter ones) had their stress on the root and others in some other syllable to the right. Stressed and unstressed vowels were the same (i.e. distinguished the same number of quantities and qualities) in Classical Latin but this changed later on. The loss of some vowel distinctions in unstressed syllables (namely between the mid vowels /ɛ/ and /e/, and between /ɔ/ and /o/) meant that some unpredictable alternations were introduced in the conjugation of some verbs. There was thus, at this stage, no way to tell from the infinitive form of verbs like \**vo'lare* ‘fly’ or \**sor'bere* ‘sip’ what vowel the rhizotonic forms of the paradigm (forms like the 2SG.PRES.IND \**volas* vs \**sorbes*) would have. Unsurprisingly if one considers Zipf’s (1935) insights, these alternations affected some of the most frequent forms of the paradigm: the singular and the 3PL forms of the present tense (of both indicative and subjunctive moods) and the singular imperative (see Table 4).<sup>3</sup>

Later phonological developments have sometimes further altered (e.g. diphthongized, see Spanish) some of these stressed vowels, and analogical changes have sometimes used the etymologically inherited stem alternation patterns as a template to distribute other formal elements (e.g. synonymous roots, see Italian) in the paradigm.

## 2.2 The L-morphome

Palatalization of consonants before front vowels or yod is such a run-of-the-mill sound change that it has happened at some point or another in the history of most languages where our knowledge of their history extends back enough in time (e.g. Slavic, English [consider ‘church’, ‘chin’ vs Dutch *kerk*, *kin*] etc.). This is true also of Romance, and, as concerns the history of the L-morphome, in at least two independent sound changes. One involved the palatalization of velars before front vowels (see *nascere* below) and the other the palatalization of non-labial consonants before /j/ (see *medir*). Because front vowels and yods were in complementary distribution in the paradigm (e.g. ‘do’: *fak-ere*, *fak-jo*, *fak-is*, *fak-it* etc.), the contexts where the two changes occurred were the exact opposite of each other and gave rise to the same pattern of stem alternation (see Table 5).

Note that the shaded cells of *nascere* are those where the sound change (e.g. *naskes*>*nastses*) did not happen whereas those of *medir* are those where the sound change (e.g. *metjo*>*metso*) did happen. This is largely inconsequential, however, as the shaded cells became, in either case, the odd-ones-out, i.e. the minority alternant

<sup>3</sup>For reasons of feature-value orthogonality (or lack thereof) and for economy of representation, the imperative will not be taken into account throughout the rest of this paper.

**Table 5** Two illustrative examples of the Romance L-morphome

	Old Spanish <i>nascer</i> ‘be born’				Portuguese <i>medir</i> ‘measure’			
	Indicative		Subjunctive		Indicative		Subjunctive	
	SG	PL	SG	PL	SG	PL	SG	PL
1	nas[k]o	nas[ts]emos	nas[k]a	nas[k]amos	meço	medimos	meça	meçamos
2	nas[ts]es	nas[ts]edes	nas[k]as	nas[k]ades	medes	medis	meças	meçais
3	nas[ts]e	nas[ts]en	nas[k]a	nas[k]an	mede	medem	meça	meçam

within their paradigms. The same as in the case of the N-morphome, analogical developments would later extend the resulting alternations to new lexemes or would use them as a model for other forms, thus revealing the productivity and cognitive reality that characterize a *bona fide* morphome.

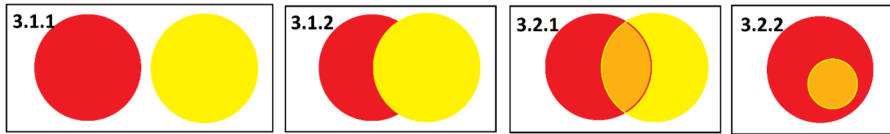
### 2.3 PYTA morphome

Somewhat differently from the two previous morphomes, PYTA (from Spanish *Preterito Y Tiempos Afines*, ‘preterite and related tenses’) does not have an origin in sound change but rather in the loss of semantic unity on a set of formerly semantically related tenses. Perfective tenses in Latin made use of a special stem (e.g. fēc- [as opposed to fac-] or rēx- [as opposed to reg-]). The common semantic thread of all those tenses, however, was progressively lost in the daughter languages so that the use of a common root for a set of unconnected tenses has become simply a morphological quirk of the languages. In Spanish, for example, the same stem (e.g. pus- < *poner* ‘put’, quis- < *querer* ‘want’, tuv- < *tener* ‘have’ etc.) is used in the past indicative, imperfect subjunctive and future subjunctive. Despite the different origin of PYTA, the same properties (i.e. analogical changes, or suppletive alternations that respect the inherited distribution) apply also to this morphome.

## 3 Logical configuration of morphomes

From a logical perspective, formal elements with a different distribution could potentially be in various different situations regarding their paradigmatic domains. To begin with, morphomes could either 3.1) constitute disjoint sets, if they do not have any cell in common, or could 3.2) overlap to some extent, when there is at least one paradigm cell that appears in both sets. Within the former we could find morphomes that 3.1.1) are completely separate and independent and 3.1.2) cases where the presence of one of the morphomes seems to entail (or even cause) the absence of the other. Within the category of morphomes that overlap we could find 3.2.1) morphomes whose distributions ‘cross-cuts’ each other and 3.2.2) morphomes which are in a subset-superset relation to each other (see Fig. 1).

All these various relations have been actually found to hold in morphomes from various (and genetically diverse) natural languages. Examples and discussion of each of these possibilities are presented next.



**Fig. 1** Various relations that may hold between morphemes

**Table 6** Extension of the formative /g/ in Spanish and Catalan, partial paradigms<sup>a</sup>

	Spanish <i>venir</i> ‘come’			Catalan <i>venir</i> ‘come’		
	PRES.IND	PRES.SUB	PAST.IND	PRES.IND	PRES.SUB	PAST.IND
1SG	vengo	venga	vine	vinc	vingui	vinguí
2SG	vienes	vengas	viniste	véns	vinguis	vingueres
3SG	viene	venga	vino	ve	vingui	vingué
1PL	venimos	vengamos	vinimos	venim	vinguem	vinguérem
2PL	veníś	vengáis	vinisteis	veniu	vingueu	vinguéreu
3PL	vienen	vengan	vinieron	vénen	vinguin	vinguerem

<sup>a</sup>Devoicing /g/>/k/ takes place in Catalan at the end of the word

### 3.1 Disjoint morphemes

Certain formal elements may sometimes be unable to appear together within the same word form. Sometimes this is merely a consequence of the fact that they appear in disjoint sets of paradigm cells. Some other times, however, forms seem to get out of their way to avoid appearing together. These two situations will be presented next.

#### 3.1.1 Morphemes that are disjoint

Sometimes, two patterns will be present in the inflectional system of a language without there being any visible interaction between the two. The frequently discussed N- and PYTA-morphemes of Romance (e.g. Maiden 2004), for example, inhabit disjoint sets of paradigm cells. As present in Spanish, the presence of one does not imply the presence of the other (e.g. *cerr-ar* ‘close’: *cierr-a* [N-morpheme] vs *cerr-ó* [no PYTA], *pon-er* ‘put’: *pon-e* [no N-morpheme] vs *pus-o* [PYTA]). In addition, the two are characterized by different forms (diphthong vs high vowel) and when both occur together in the same lexeme (e.g. *quer-er* ‘want’: *quier-e* [N-morpheme] vs *quis-o* [PYTA], *pod-er* ‘be able to’: *pued-e* [N-morpheme] vs *pud-o* [PYTA]) there is no change in either the form or the distribution of either of the morphemes.

It might seem that lack of morphosyntactic overlap could prevent morphemes from influencing each other, however, this turns out to be incorrect. The L-morpheme (see the distribution of the alternant *veng-* below) and PYTA (illustrated in Table 6 by the past indicative) are also similarly independent in Spanish. In closely-related Catalan, however, we can see how the two interact to some extent (see Table 6).

The velar consonant that usually characterizes the L-morpheme roots (e.g. see Spanish) has extended in Catalan to the domain of PYTA (see Wheeler 2011), thus

**Table 7** Present paradigms of two Spanish verbs showing stem alternation

	<i>entender</i> ‘understand’			<i>poner</i> ‘put’	
	Indicative	Subjunctive		Indicative	Subjunctive
1SG	entiendo	entienda		pongo	ponga
2SG	entiendes	entiendas		pones	pongas
3SG	entiende	entienda		pone	ponga
1PL	entendemos	entendamos		ponemos	pongamos
2PL	entendéis	entendáis		ponéis	pongáis
3PL	entienden	entiendan		ponen	pongan

**Table 8** Paradigm of Spanish *venir*

	Indicative	Subjunctive
1SG	vengo	venga
2SG	vienes	vengas
3SG	viene	venga
1PL	venimos	vengamos
2PL	venís	vengáis
3PL	vienen	vengan

bestowing a certain unity upon erstwhile independent stem alternants. This is just meant to remind us that morphemes can interact and influence each other even in the cases when they in principle have disjoint distributions.

### 3.1.2 Morphomes which exclude each other

Some cases of non-overlap between morphemes, however, are quite different, in that a lack of overlap seems to be an “actively sought” state of affairs in the language. In some cases, thus, even when the forms and patterns involved would be theoretically compatible within a single word form (as witnessed by the distribution of the relevant forms when they occur in isolation) some morphemes seem to prevent the application of the other instead. Consider the behaviour of the reknown N- and L-morphemes in Spanish (see Table 7).

The paradigmatic distribution of the N-morphome (illustrated above with the verb *entender*) and that of the L-morphome (illustrated above with the verb *poner*) overlap in their morphosyntactic domain and involve different forms. Accordingly, they should theoretically be able to appear together in the paradigm cells where both apply (i.e. 1SG.IND, SG.SUB and 3PL.SUB). Instead of cross-cutting one another in these situations, however, the L-morphome appears to ‘block’ the N-morphome in Spanish. Consider the paradigm below (see Table 8).

In the paradigm of Spanish *venir* ‘come’ (also of *tener* ‘have’), diphthongization does not occur throughout the rhizotonic cells as is usually the case. Instead, diphthongization is restricted to those rhizotonic cells that are not themselves part of the

**Table 9** Verb ‘come’ in two diachronic stages of Romance, reconstructed forms

	‘come’ in Proto-Romance		‘come’ in Proto-Ibero-Romance	
	Indicative	Subjunctive	Indicative	Subjunctive
1SG	<i>‘venjo</i>	<i>‘venja</i>	<i>‘vejo</i>	<i>‘veja</i>
2SG	<i>‘venes</i>	<i>‘venjas</i>	<i>‘venes</i>	<i>‘vejas</i>
3SG	<i>‘venet</i>	<i>‘venjat</i>	<i>‘vene</i>	<i>‘veja</i>
1PL	<i>ve’nimus</i>	<i>ve’njamus</i>	<i>ve’nimos</i>	<i>vi’jamos</i>
2PL	<i>ve’nites</i>	<i>ve’njates</i>	<i>ve’nites</i>	<i>vi’jates</i>
3PL	<i>‘venent</i>	<i>‘venjant</i>	<i>‘venen</i>	<i>‘vejan</i>

**Table 10** Stem vowel distribution inherited from the effects of rhizotony and yod in ‘sleep’ (left) and modern Spanish and Portuguese reshapings (adapted from O’Neill 2012)

	Proto-Ibero-Romance		Spanish		Portuguese	
	SG	PL	SG	PL	SG	PL
1SG	<i>*dormo</i>	<i>*dorma</i>	<i>duermo</i>	<i>duerma</i>	<i>durmo</i>	<i>durma</i>
2SG	<i>*dormes</i>	<i>*dormas</i>	<i>duermes</i>	<i>duermas</i>	<i>dormes</i>	<i>durmas</i>
3SG	<i>*dorme</i>	<i>*dorma</i>	<i>duerme</i>	<i>duerma</i>	<i>dorme</i>	<i>durma</i>
1PL	<i>*dormimos</i>	<i>*durmamos</i>	<i>dormimos</i>	<i>durmamos</i>	<i>dormimos</i>	<i>durmamos</i>
2PL	<i>*dormites</i>	<i>*durmates</i>	<i>dormís</i>	<i>durmaís</i>	<i>dormis</i>	<i>durmais</i>
3PL	<i>*dormen</i>	<i>*dorman</i>	<i>duermen</i>	<i>duerman</i>	<i>dormem</i>	<i>durmam</i>

domain of the L-morpheme (instantiated by its exponent /g/ in this case). Thus, the presence of the stem extension /g/ (characteristic of the L-morpheme in Spanish) seems to inhibit the diphthongization typical of the N-morpheme.

A widespread interpretation of this apparent blocking is that it is the straightforward result of sound changes from Latin and that, accordingly, there is nothing extraordinary about it really. Because the yod that created the consonant stem alternations typical of the L-morpheme also had a raising effect on previous mid vowels in various Iberian Romance varieties, it sometimes ‘bled’ the later diphthongization (/e/ > /je/) of the N-morpheme in a natural, systemic way (see Table 9).

However, and apart from the restoration of stem vowel identity within the L-morpheme in the daughter languages (i.e. *vi’jamos* > *ve’jamos*, *vi’jates* > *ve’jates*), there are a few things to note in relation to this. The first is that this ‘bleeding’ that restricted diphthongs from the N-morpheme to a subset of it must have happened with many other lexical items too (e.g. *sentir* ‘feel’ or *dormir* ‘sleep’, see Table 10). Yet, in contemporary Spanish, this sound change has been generally turned back unless an L-morpheme exponent also appears. Another interesting point is that, although it is usually believed (Penny 2002:174) that /j/ did not last long enough in the -er conjugation to produce metaphonic effects, the same pattern that appears in *venir* also appears in *tener* ‘have’. In addition, a number of verbs that in Old Spanish had both diphthong and velar augment (e.g. *doler* ‘hurt’, *duele* [N-morpheme], *duelga* [L-morpheme & N-morpheme]) have since lost the latter. The result is that diphthongization in Span-



ish conjugation is only ever restricted to 2/3SG and 3PL indicative nowadays in the presence of an overt exponent for the L-morphome. In the opposite direction, the implicature also holds: Diphthongization never happens, in Spanish conjugation, in the presence of an overt L-morphome.

Although different forms (e.g. stem vowel and stem extension) may occasionally exclude one another, the compatibility of two cross-cutting patterns is likely to be particularly difficult, of course, when they involve the same locus in the word. As I have just mentioned, in some Iberian Romance varieties the L-morphome had vocalic in addition to consonantal instantiations, which put it in direct competition with the N-morphome. In some lexemes, both patterns should have been simultaneously present in the vowel stem. However, when coexisting in a single lexical item, one pattern necessarily implies the disruption of the other. This circumstance may be undesirable for language users' prediction and acquisition of these patterns, which may be the reason why these conflicts tended to be resolved, rather than perpetuated, by making these aberrant, hybrid patterns conform to one of the more widespread and simple ones (see Table 10).

In Table 10 above it is shown how the simultaneous application of the sound changes that led to the Romance N and L-morphomes would have resulted, in *dormir* 'sleep' and other verbs, in a pattern of considerable complexity and in a disruption to the usual stem affinities inherent to the N and L-morphomes. Probably as a consequence, the inherited pattern has been analogically modified both in Spanish and in Portuguese but in diverging ways. Spanish has proceeded by generalizing the diphthong throughout the cells of the N-morphome. Portuguese, by contrast, has preferred to preserve the integrity of the L-morphome by generalizing the high vowel to all of its cells. Whether or not these synchronic and diachronic facts should be interpreted along the lines of one morphome 'blocking' another is open to dispute and beyond the purposes of this paper.

### 3.2 Morphomes in non-disjoint relations

In contrast to the cases presented in the previous section, some formal elements and distributions do not appear to be incompatible. Different forms, thus, may appear in these cases together in one and the same word form. This may occur regardless of whether the distribution of one of them is a subset of the other or not.

#### 3.2.1 Morphomes which cross-cut each other

In contrast with the cases presented in the previous section, the N- and L-morphomes can some other times be fully compatible. Consider the following paradigms (see Table 11).

In remarkable contradistinction to the paradigm of its Spanish cognate (Table 8), Ansotano Aragonese 'come' does not restrict its diphthongized stem alternant to those cells outside of the L-morphome (Barcos 2007). The N- and L-morphomes are also compatible in Portuguese, as witnessed by the cross-cutting distributions of their respective exponents /ε/ and /k/ in the verb 'lose'.

**Table 11** Cross-cutting, compatible morphomic exponents

	Ansotano Aragonese 'have' (Barcos 2007)		Portuguese 'lose'	
	IND	SUB	IND	SUB
1SG	bjengo	bjengaj	p[ɛ]rco	p[ɛ]rca
2SG	bjen(e)s	bjengas	p[ɛ]rdes	p[ɛ]rcas
3SG	bjene	bjenga	p[ɛ]rde	p[ɛ]rca
1PL	be'nimos	ben'gamos	perdemos	percamos
2PL	be'niθ	ben'gaθ	perdeis	percais
3PL	bjenen	bjengan	p[ɛ]rdem	p[ɛ]rcam

**Table 12** Pite Sami *basset* 'fry' (Wilbur 2015: 174)

	PRES				PAST		
	SG	DU	PL		SG	DU	PL
1	bas-áv	biss-in	bass-ep		biss-iv	bis-ijmen	bis-ijmä
2	bas-á	bass-ebähten	bass-ebähtet		biss-e	bis-ijden	bis-ijdä
3	bass-a	bass-eba	biss-e		bis-ij	bis-ijga	biss-in

Although the literature on morphemes and morpheme clashes has been dominated by evidence from Romance, this is by no means a phenomenon exclusive to that language family. Some other highly inflecting languages (particularly of the fusional type) show cross-cutting morphological phenomena that should be considered morphomic because they are synchronically not phonological and apply to an unnatural set of morphosyntactic contexts. This is the case, for example, of consonant gradation and vowel apophony in Sami, see Table 12

As explained by Wilbur (2015), consonant gradation (strong grade indicated in light grey above) is no longer phonologically predictable in Pite Sami (nor in the other Sami varieties for that matter). Also, although Wilbur (2015) chooses the term “vowel harmony” to refer to vowel alternations like the one shaded in dark gray here, these are no longer phonologically determined assimilations in Pite Sami. As a result we have that in paradigms like that of *basset* ‘fry’ above, the two morphological processes cross-cut, as the L- and N-morphemes in Table 11, to generate a total of four different stem alternants. Two morphomic forms, however, can be compatible without this resulting in a 4-way division. This is what happens with morphological elements in subset-superset relations, a circumstance that is discussed in the next section.

### 3.2.2 Morphemes in subset relations to each other

Unnatural patterns of morphological identity in one language can also be in subset relations to each other. This is the case, for example, of the syncretisms found in the inflecting verbs of Skou (Donohue 2004:219), see Table 13.

As shown in Table 13, there are in Skou several progressively larger unnatural sets of paradigm cells that are syncretic. The conflated cells highlighted for the verbs

**Table 13** Exponence patterns of various Skou verbs

	2PL	1SG	3SG.NF	1PL	3PL	3SG.F	2SG
eat	ang	kang	kang	nang	tang	kang	mang
walk	ha	ha	ka	na	ta	wa	ma
narrate	láng	láng	láng	táng	jáng	wáng	páng
hit	ká	ká	ká	ká	já	wá	bá
utter	lá	lá	lá	lá	lá	wá	pá
get.F	wé	wé	wé	wé	wé	wé	pé
plant	wá	wá	wá	wá	wá	wá	wá

**Table 14** Stem form of the verbs ‘listen’ and ‘spill’ in Lealao Chinantec (Rupp 1996: 424, 427–429)

	listen				spill		
	PRES	FUT	PAST		PRES	FUT	PAST
1SG	nuu	nuu	nuu		tũũ	tũũ	tũũ
1PL INC	nuu	niuu	niuu		tiũũ	tiũũ	tiũũ
1PL EX	nuu	niuu	niuu		tiũũ	tiũũ	tiũũ
2SG	nuu	nuu	niuu		tũũ	tũũ	tiũũ
2PL	nuu	nuu	niuu		tũũ	tũũ	tiũũ
3	nuu	nuu	nuu		tiũũ	tiũũ	tiũũ

‘walk’, ‘narrate’, ‘hit’, ‘utter’ and ‘get.F’ are all unnatural sets of cells which in some lexemes share a form to the exclusion of the neighbouring paradigm cells. These patterns of morphological identity are, like Russian matryoshkas, progressively smaller subsets of each other.

Crucially, however, morphemes can be nested into one another not only when comparing formal patterns found in different lexical items but also (when different forms are involved of course, see Sect. 3.3) within the paradigm of a single lexeme. Consider the distribution of stem alternants in the following Lealao Chinantec verbs (cited in Baerman and Corbett 2012, after Rupp 1996) (see Table 14).

The stem *niuu* appears, as shown in Table 14, in a subset of the paradigm cells that the stem *tiũũ* applies to. The two stem alternation patterns can appear independently, as illustrated above, but can also appear simultaneously, as observed in the paradigm of the verb ‘take’, see Table 15.

A suppletive stem alternation (*uuyh* vs *hṽh*) can be found with the same distribution as *tũũ* vs *tiũũ* before. Within the part of the paradigm that hosts the alternant *hṽh*, however, a vowel apophony (*hih* vs *hiḥ*) distinguishes a smaller subset of cells that is identical to that of *niuu* before. We can see, therefore, that morphomic elements can sometimes coexist in subset relations to each other.

**Table 15** Stem form of ‘take’ in Lealao Chinantec (Rupp 1996:460)

	‘take’		
	PRES	FUT	PAST
1SG	uuyh	uuyh	uuyh
1PL INC	hĩh	hĩh	hĩh
1PL EX	hĩh	hĩh	hĩh
2SG	uuyh	uuyh	hĩh
2PL	uuyh	uuyh	hĩh
3	hĩh	hĩh	hĩh

**Table 16** Three Luxembourgish verbs, present tense (Wunderlich 2012:171 after Nübling 2000)

	‘be’		‘have’		‘go’	
	SG	PL	SG	PL	SG	PL
1	si-n	si-n	hu-n	hu-n	gi-n	gi-n
2	bas	si-d	hues	hu-t	gees	gi-t
3	as	si-n	huet	hu-n	geet	gi-n

### 3.3 A note on compatibility

In order to spot patterns that overlap within a single lexeme’s paradigm, the patterns have to be, obviously, instantiated by different formal elements. This is the case of the nested patterns displayed by the stem of Lealao Chinantec ‘take’ above, since one pattern involves full suppletion while the other makes use of vowel apophony in the stem. When different structural elements are involved (e.g. stem vs affix, or affix1 vs affix2) this may be more likely to result in patterns which are compatible within a given lexeme’s paradigm within a single word form. Consider, for example, the following Luxembourgish verbs, see Table 16.

The morphosyntactic distribution of the suffix *-n* as well as that of stem alternants *si-*, *hu-* and *gi-* is morphosyntactically unnatural. The set of contexts for *-n* (1SG, 1PL, 3PL) is a subset of those of the stem alternants *si-*, *hu-* and *gi-* (1SG, 1PL, 3PL, 2PL). Since the exponents are different, the patterns are compatible in principle and thus it is an empirically decidable fact whether they actually are.

It has to be kept in mind, however, that compatibility becomes a logical impossibility when two patterns are marked in the same place in the word. This is the case, for example, of raising (e.g. *pedir* ‘ask for’) and diphthongization (e.g. *segar* ‘reap’) in Spanish verb conjugation (see Tables 17 and 18).

Because the forms that instantiate these two patterns are incompatible (i.e. the vowel nucleus of a stem cannot be simultaneously /je/ and /i/), one necessarily has to exclude the other (see Table 19).

Because of this, blocking is the only possibility in these situations. If the two forms are found within a single lexeme, as in *sentir* above, only the more restricted one can possibly block the other. That is, only the subset can block the superset. Even though the restrictive properties and empirical status of Pāṇinian blocking are beyond the scope of the present paper this is a point to keep in mind.

**Table 17** Distribution of high vowel stem in Spanish rising verbs

	PRES.IND	PRES.SUB	IPF	PAST	IPF.SUB I	IPF.SUB II	FUT	COND
1SG	pido	pidá	pedía	pedí	pidiera	pidiese	pediré	pediría
2SG	pides	pidas	pedías	pediste	pidieras	pidieses	pedirás	pedirías
3SG	pide	pidá	pedía	pidió	pidiera	pidiese	pedirá	pediría
1PL	pedimos	pidamos	pedíamos	pedimos	pidiéramos	pidiésemos	pediremos	pediríamos
2PL	pedís	pidáis	pedíais	pedisteis	pidierais	pidieseis	pediréis	pediríais
3PL	piden	pidan	pedían	pidieron	pidieran	pidiesen	pedirán	pedirían

**Table 18** Distribution of diphthong stem in Spanish N-morphome verbs

	PRES.IND	PRES.SUB	IPF	PAST	IPF.SUB I	IPF.SUB II	FUT	COND
1SG	siego	siegue	segaba	segué	segara	segase	segaré	segaría
2SG	siegas	siegues	segabas	segaste	segaras	segases	segarás	segarías
3SG	siega	siegue	segaba	segó	segara	segase	segará	segaría
1PL	segamos	seguemos	segábamos	segamos	segáramos	segásemos	segaremos	segaríamos
2PL	segáis	seguéis	segabais	segasteis	segarais	segaseis	segaréis	segaríais
3PL	siegan	sieguen	segaban	segaron	segaran	segasen	segarán	segarían

**Table 19** Cooccurrence of diphthong and high vowel stem in *sentir* ‘feel’

	PRES.IND	PRES.SUB	IPF	PAST	IPF.SUB I	IPF.SUB II	FUT	COND
1SG	siento	sienta	sentía	sentí	sintiera	sintiese	sentiré	sentiría
2SG	sientes	sientas	sentías	sentiste	sintieras	sintieses	sentirás	sentirías
3SG	siente	sienta	sentía	sintió	sintiera	sintiese	sentirá	sentiría
1PL	sentimos	sintamos	sentíamos	sentimos	sintiéramos	sintiésemos	sentiremos	sentiríamos
2PL	sentís	sintáis	sentíais	sentisteis	sintierais	sintieseis	sentiréis	sentiríais
3PL	sienten	sintan	sentían	sintieron	sintieran	sintiesen	sentirán	sentirían

## 4 Interactions between morphemes

When they “inhabit” the same inflectional system in a language, two or more morphemes may coexist (unperturbed), as in some of the examples that were presented in the previous section (e.g. Tables 11 and 15) or they may instead interact with each other or even “conspire” in various ways to give birth to new morphemes. The latter is likely to be especially frequent when they comprise overlapping sets of cells or when they cooccur within a single lexical item.

It has to be kept in mind that the motivation for this tendency is not self-evident. Formal elements with a different paradigmatic distribution are usually, and canonically (Corbett 2007), orthogonal. When the distributions of two formal elements

**Table 20** Separative (left) and cumulative (right) fictitious paradigms

	Paradigm A			Paradigm B		
	SG	DU	PL	SG	DU	PL
1	-ga-lo	-ga-ze	-ga-fy	-moli	-fodo	-pemu
2	-mu-lo	-mu-ze	-mu-fy	-tura	-syge	-hovy
3	-si-lo	-si-ze	-si-fy	-hane	-luba	-tiro

**Table 21** Tiwi agreement morphology (Lee 1987: 173)

	Prefixes (first part)			Prefixes (second part)			Prefixes (combined)		
	SG	DU	PL	SG	DU	PL	SG	DU	PL
1EX	ngi-	ngi-	ngi-	-Ø-	-mpi-	-mpi-	ngi-	ngimpi-	ngimpi-
1IN	–	mu-	nga-	–	-Ø-	-Ø-		mu-	nga-
2	nyi-	ngi-	ngi-	-mpi-	-mpi-	-mpi-	nyimpi-	ngimpi-	ngimpi-
3.F	a-	wu-	wu-	-mpi-	-Ø-	-Ø-	ampi-	wu-	wu-
3.M	a-	wu-	wu-	-Ø-	-Ø-	-Ø-	a-	wu-	wu-

cross-cut, this constitutes a trait that may well be advantageous to language-users, as this permits to multiply morphosyntactic distinctions with a more reduced inventory of forms. Compare these two paradigms in Table 20.

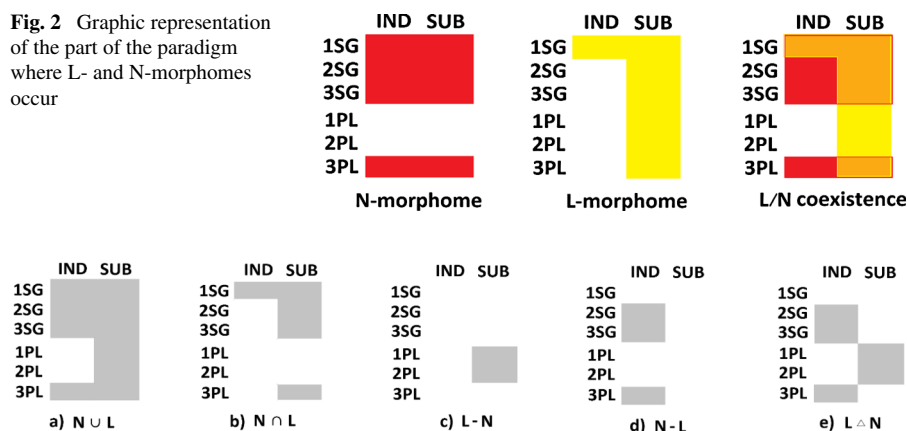
The two paradigms make the same number of distinctions. However, if we were to enumerate the exponence rules for the two systems we would have fewer (6) and simpler (one value) entries in the separative paradigm (e.g. 1 > ga) compared to the cumulative one (which would need 9 rules, each with 2 value-specifications [e.g. 1.SG > -moli]). The economy brought forth by a smaller number of forms that cross-cut to maximize morphosyntactic distinctions also holds with forms with an unnatural morphosyntactic distribution when these cross-cut each other. Consider the following paradigm, see Table 21.

Because the two formatives highlighted above cross-cut each other rather than constitute disjoint sets or subset-supersets of each other, the two forms are able to achieve three rather than two formal distinctions. It is, thus, not self-evident that cross-cutting, orthogonal morphemes like the ones in Table 11 would be undesirable. Moreover, two or more forms with unnatural morpho-syntactic distributions could in principle combine into natural morphosyntactic distinctions, as happens in Tiwi to some extent.

It must be kept in mind, however, that the functional load of morphomic elements is sometimes very low. Stem alternations in the most conservative Romance varieties, for example, are almost 100% redundant,<sup>4</sup> since the relevant distinctions are usually expressed by the affixes. Some functional concerns, thus, may have little value to

<sup>4</sup>The only place I can think of where stem alternations can introduce additional formal distinctions in Spanish are cases where the regular whole-word syncretism between 1PL present and past (e.g. amamos/amamos, vivimos/vivimos) is avoided because of the use of a stem alternant in the past (e.g. estamos/estuvimos, decimos/dijimos).

**Fig. 2** Graphic representation of the part of the paradigm where L- and N-morphemes occur



**Fig. 3** Possible interactions between the N and L-morphemes

predict the behaviour of morphemes and, thus, their interaction should definitely be investigated by engaging with the empirical facts.

Because of the sheer amount of data available in the linguistic literature for the morphemes of Romance compared to those of other language families I will focus here on the interactions between those morphemes. More concretely, concentrating on the interactions between the N- and the L-morphome is especially convenient because they constitute overlapping sets of cells and often affect the same lexemes. The schematic representation of the paradigmatic distribution of these two morphemes will follow here the traditional convention of representing person-number values non-orthogonally, which makes it possible to display these in a two-dimensional format (see Fig. 2).

Based on the basic operations for constructing new sets from given sets in Set Theory we can elaborate another pre-empirical typology of morphome interactions. Two morphemes could conspire to create new morphological patterns in the following ways:

- Union ( $A \cup B$ ): the set of cells included in either one of the morphemes or in both.
- Intersection ( $A \cap B$ ): the set of cells included in both morphemes simultaneously.
- Relative complement of A in B ( $B - A$ ): the set of cells included in morphome B but not in A.
- Relative complement of B in A ( $A - B$ ): the set of cells included in morphome A but not in B.
- Symmetric difference ( $A \Delta B$ ): the set of cells in either one of the morphemes but not in both.

Here is a graphic representation of what each of these sets would include when applied to the N- and L-morphemes of Romance, see Fig. 3.

It is, I believe, extraordinary, that most of these patterns are actually found somewhere in Romance, not simply as the accidental product of sound change but as a target category of analogical morphological innovations. This is shown in the following sections.

**Table 22** N U L distributions in some Romance varieties

	Verb ‘have’ in Old French		Verb ‘measure’ in Spanish		Verb ‘have to’ in Savognin (Maiden 2018:213)	
	Indicative	Subjunctive	Indicative	Subjunctive	Indicative	Subjunctive
1SG	tieng	tiegne	mido	mida	stó	stóptga
2SG	tiens	tiegues	mides	midas	stóst	stóptgas
3SG	tient	tiegne	mide	mida	stó	stóptga
1PL	tenons	tiegniens	medimos	midamos	duágn	stóptgan
2PL	tenez	tiegniez	medis	midais	duéz	stóptgas
3PL	tient	tiegnent	miden	midan	stón	stóptgan

#### 4.1 Union (N U L)

The analogical impulse to spread particular forms or alternants to the whole of N U L is a well-known tendency in Romance. This can be implemented with diverse forms (see Table 22).

Consider the Old French paradigm above. As a result of sound change, stem vowel diphthongization (i.e. /e/ > /je/) and palatalization of the last consonant of the stem (i.e. /n/ > /ɲ/) <sup>5</sup> should be characteristic of the N-morphome and L-morphome cells respectively and should thus cross-cut each other in Old French. However, the vowel within the NUL paradigm cells has been levelled analogically in favour of the diphthong, which after the change opposes NUL to the rest of the paradigm. The diphthong did not spread beyond the NUL cells which thus acted as a niche (Aronoff 2016) for that particular morphological element.

In the Old French case, the formal element characteristic of the N-morphome (i.e. a diphthong) is generalized to the whole of NUL. Something quite different happens in the Spanish case above. Rising (i.e. /e/ > /i/) in Spanish is the result of anticipatory assimilation of mid vowels before a yod (i.e. \*metjo > mido, \*metimus > medimos). Sound changes in the presence of this yod are precisely what gave birth the L-morphome in Romance (see Table 9) and thus, the rising would have occurred, initially, in just those cells. We can see, however, that in Spanish, like in Old French, a single vowel has been generalized to the whole of NUL. In this case, by contrast, it is the vowel that originally characterized the L-morphome.

Another striking example of how NUL can act as a morphological class in Romance is the paradigm of Savognin *duéir* ‘have to’. As explained by Maiden (2018:213), in this verb, suppletion <sup>6</sup> occurs in the set of cells defined by the union of the N- and the L-morphomes. A specific root alternant occurs in these contexts in the paradigms of other lexemes and this fact provides a niche, model, or template for the paradigmatic distribution of other formal elements.

<sup>5</sup>Both “gn” and “ng” are orthographic representations of /ɲ/.

<sup>6</sup>It is not entirely clear whether one should describe this as suppletion or as defectivity of the verb *duéir*. The shaded cells of the paradigm are (taken) from another verb (*stueir*) where they are not suppletive. See Maiden (2018:213).



**Table 23** Alta Ribagorza Aragonese (O'Neill 2018, from Haensch 1958:157–186)

	'have'		'twist'		'sleep'	
	IND	SUB	IND	SUB	IND	SUB
1SG	tjéngo	tjénga	twérsko	twérska	dwérmo	dwérma
2SG	tjéns	tjéngas	twérses	twérskas	dwérmes	dwérmas
3SG	tjéne	tjénga	twérse	twérska	dwérme	dwérma
1PL	tením	tjéngám	torsém	twerskám	dormím	dormám
2PL	teníts	tjéngáts	torséts	twerskáts	dormíts	dormáts
3PL	tjénen	tjéngan	twérsen	twérskan	dwérmen	dwérman

#### 4.1.1 A note on 'interaction'

It has been claimed more or less explicitly throughout this section that these are new patterns that arise from the **interaction** between the two morphemes. However, finding analogical processes that affect or result in distributions like NUL is arguably not incontrovertible evidence that this is due to the interaction between the N and L-morphemes. It is far from inconceivable, for example, that the extension in Old French of diphthongization to 1PL and 2PL subjunctive had nothing to do with the L-morphome *per se* and constituted simply a formal levelling within the present subjunctive. Is there a way to tell?

I believe sometimes there is. If a given form's distribution has originated from the interaction between the two morphemes, one would logically expect to find, at least occasionally, that the form has a different distribution in the presence of an overt marking of the other morphome and in the absence of it. As a matter of fact we do come across such cases. Consider the following described in Table 23.

The case is similar to the interaction between diphthongization and the augment (i.e. stem-final /g/ or /k/) described for Spanish in Table 8. The morphosyntactic extent of diphthongization in any particular verb is dependent on whether the L-morphome is present or not.<sup>7</sup> The (notable) difference is that, in Alta Ribagorza Aragonese, the L-morphome augment necessitates diphthongization rather than rejects it. As a consequence, one morphome becomes a subset of the other one.

Deep down, however, both strategies may be just different solutions to one and the same problem: the orthogonality of morphomic forms and the multiplication of stem alternants that this originates. By making diphthongization and the velar augment mutually exclusive, Spanish reduces from four to three the number of stem alternants in the lexemes where the two morphemes cooccur. Alta Ribagorza Aragonese, paradoxically, achieves the exact same effect by making diphthongization a necessary concomitant of the velar augment.

<sup>7</sup>In an additional taxonomization of morphomic patterns one could distinguish "dependent" from "independent" morphemes on the basis of whether a particular pattern can occur on its own or not. Spanish diphthongization in 2/3SG and 3PL present indicative and Alta Ribagorza Aragonese diphthongization in SG, 3PL and 1/2PL.SUB would count, thus, as dependent morphemes, since they never occur in the absence of an overt L-morphome.

**Table 24**  $N \cap L$  distributions in some Romance varieties (Maiden 2012)

	Pignia Romansh ‘do’		Bolognese ‘go’		Genoese ‘give’	
	IND	SUB	IND	SUB	IND	SUB
1SG	fēc	fēci	va:g	va:ga	dagu	dage
2SG	fas	fēcas	vε	va:g	de	dagi
3SG	fa	fēci	va	va:ga	da	dage
1PL	fazain	fazejan	andain	andannja	demu	demu
2PL	fazets	fazejas	ande	andēdi	dε	dε
3PL	fan	fēcan	van	va:gen	dan	dagan

**Table 25**  $L-N$  distributions (Maiden 2012)

	Felechosa Asturian ‘do’		Felechosa Asturian ‘bring’	
	IND	SUB	IND	SUB
1SG	fago	faga	trao	traa
2SG	faes	fagas	traes	traas
3SG	fai	faga	trae	traa
1PL	faθemos	fifamos	traemos	trifamos
2PL	faθeis	fifaos	traes	trifaos
3PL	faen	fagan	traen	traan

## 4.2 Intersection ( $N \cap L$ )

Some other times, interaction between the N- and L-morphemes results in the intersection of the two being characterized by a particular form, see Table 24.

In the paradigms above we see how  $N \cap L$  cells share a common stem to the exclusion of the rest. This situation originated analogically when the original L-morpheme forms were evicted from the 1PL and 2PL subjunctive, sometimes, but not always resulting in whole-word syncretism with the indicative (e.g. Genoese).

## 4.3 Relative complement of N in L ( $L - N$ )

Occasionally, the relative complement of N in L (i.e. the 1PL and 2PL cells of the present subjunctive) can also act *en bloc* in processes of analogical change, see Table 25.

In Felechosa Asturian, for example, as presented in Maiden (2012), the stem alternant that in other Romance varieties characterizes the preterite and related tenses (i.e. the PYTA root) has been extended to  $L - N$ , thus giving rise to a contrast with the rest of the present paradigm.

## 4.4 Relative complement of L in N ( $N - L$ )

Some other times, it is the cells of the N-morpheme that do not also form part of the L-morpheme that have forms of their own (see Table 26 and Table 8 as well).

**Table 26** N–L distributions in some Romance varieties

	Portuguese ‘run’		Lags Romansh ‘let, cause’ (Maiden 2018:108)		Palmoli ‘give’ (Maiden 2018:208)
	IND	SUB	IND	SUB	IND
1SG	c[o]rro	c[o]rra	lafel	lafi	ðiaŋgə
2SG	c[ə]rres	c[o]rras	lais	lafies	ðellə
3SG	c[ə]rre	c[o]rra	lai	lafi	ðallə
1PL	corremos	corramos	fein	fejen	ðajmə
2PL	correis	corrais	feis	fejes	ðat
3PL	c[ə]rrem	c[o]rram	lain	lafien	ðallə

The Portuguese case shows the analogical extension of a vowel alternation pattern affecting N–L. The verb *correr* (from Latin *currō*) should not have any vowel or consonant alternations etymologically (consider Spanish *corro corres corre corremos corréis corren*). However, it was analogically introduced in Portuguese *correr* (also in others like *beber*, see Maiden 2012) on the basis of other verbs that would regularly have had the alternation.

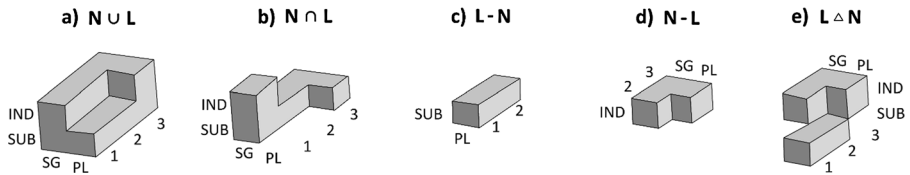
The paradigm of Lags Romansh shows an expected root (i.e. containing /f/ from Latin *laxō*) in the unshaded cells. A root without a stem-final consonant is used, by contrast, in N–L. As reported in Maiden (2018:108), this variant may have been innovated initially for the imperative on the basis of other verbs and extended to the cells shaded above.

Last, the paradigm from the Romance variety of Palmoli shows the use of a formal element -ll- in those same paradigm cells. As explained by Maiden (2018: 208) this probably originated also in imperative contexts by the reanalysis of what was originally an object clitic as part of the root. The singular imperative, similarly to the paradigm cells that constitute N–L, is rhizotonic and does not participate in the yod-related changes of the L-morphome. Because of this morphological affinity, the developments of Lags Romansh and Palmoli Romance “make sense”. After all, N–L is the smallest morphomic niche to which forms originating in the imperative may spread.

#### 4.5 Symmetric difference (L $\Delta$ N)

The last set that was presented in the range of logical possibilities for morphome interactions involved what Set Theory refers to as ‘symmetric difference’. This would arguably be a more complex interaction between morphomes than the earlier ones in that it constitutes the sum of previous categories (i.e.  $[N - L] + [L - N]$ ) or their subtraction (i.e.  $[N \cup L] - [N \cap L]$ ). Because of this, it may not necessarily be regarded as unexpected that no analogical developments have been found where this set of cells has behaved as a unit. In addition, unlike every other set discussed here, the symmetric difference of the N- and L-morphomes constitutes a geometrically discontinuous swath of the paradigm, see Fig. 4.

This could turn out to be a relevant property distinguishing L $\Delta$ N from the other sets analyzed here (see Pertsova 2011 and McCreight and Chvany 1991). It might be



**Fig. 4** Morphosyntactic geometry of the earlier sets of paradigm cells

**Table 27** Stem vowel distribution inherited from the effects of rhizotony and yod (left) in ‘sleep’ and modern Spanish and Portuguese reshapings (adapted from O’Neill 2012)

	Proto-Ibero-Romance		Spanish		Portuguese	
	SG	PL	SG	PL	SG	PL
1SG	*dormo	*dorma	duermo	duerma	durmo	durma
2SG	*dormes	*dormas	duermes	duermas	dormes	durmas
3SG	*dorme	*dorma	duerme	duerma	dorme	durma
1PL	*dormimos	*durmamos	dormimos	durmamos	dormimos	durmamos
2PL	*dormites	*durmates	dormis	durmais	dormis	durmais
3PL	*dormen	*dorman	duermen	duerman	dormem	durmam

cognitively more complicated for 2/3SG.IND, 3PL.IND and 1/2PL.SUB to constitute the domain for a grammatical generalization of any sort, which is what is required for a morpheme to arise. The greater unnaturalness of this set of cells compared to the others is also reflected on its morphosyntactic coherence (Esher 2014).<sup>8</sup> The symmetric difference of L and N ranks lowest among these sets according to this measure (26.6%), followed by N–L (33.3%), N∪L (42.8%), N∩L (46%) and L–N (50%).

The non-existence of LΔN as a productive morphomic category in Romance is empirically especially relevant because it can be argued that language users have sometimes let golden opportunities pass by in this respect. For example, when the common ancestor of Spanish and Portuguese acquired by sound change a pattern of vocalic stem alternation that singled out the forms in LΔN from those in the rest of the paradigm (see Table 27), an analogical levelling could well have matched the vowels in N–L and L–N. Imagine, for example, an analogical reshaping of 1PL.SUB and 2PL.SUB on the basis of the 3SG.IND (i.e. *dur'mamos*>*dormamos*, *dur'mates*>*dormates*). None of the daughter varieties have proceeded this way, as far as I know, and analogical processes have aimed, instead, at restoring the formal identities within N (Spanish) or L (Portuguese) (see Table 27).

Table 27 above, as well as the evidence presented in previous sections, therefore, shows a possible constraint on the conflatability of stem spaces (Montermini and Bonami 2013) arising from the combination of orthogonal morphomic distinctions. Every one of the stem spaces that arise from this (see N∩L, L–N and N–L) can

<sup>8</sup>This is the average number of feature values shared between the cells that make up some subset of the paradigm.

constitute the domain for later analogical changes, which shows that they can be cognitively active morphological categories. In addition to this, all their combinations except for one can act in the same way. The Spanish development in Table 27 shows the conflation of  $N \cap L$  and  $N-L$ , which amounts to restoring the formal unity of the N-morphome. The Portuguese analogical change, in turn, shows conflation of  $N \cap L$  and  $L-N$ , which amounts to a restoration of the formal unity of the L-morphome. We also saw (Table 22) that  $N \cap L$ ,  $N-L$  and  $L-N$  can also act together as a unit in analogical change and appear to be, thus, also “mergeable” in a single stem space or morphome.

The only combination of stem spaces that has not been found here to be a possible domain for analogical innovations and thus for a well-defined morphomic category is that of  $N-L$  and  $L-N$  (i.e.  $L \Delta N$ ). If the impossibility of such a morphomic category holds across Romance varieties and if that category ( $A \Delta B$  more generally) is not found across languages either (or if it is found to be exceedingly infrequent), this would constitute a significant constraint on the possibilities of morphomic architecture and a window into the cognitive representations of language users. More specifically, this would remind of the so-called \*ABA constraints in morphological exponence (e.g. Bobaljik and Sauerland 2018) whereby a given morphological element can serve as an expression of more than one category only when these are “adjacent” within some ordered scale. The difference, of course, is that these categories are here morphomic and inherently meaningless. It is unclear, however, whether such a factor is of any relevance to language users.

## 5 Conclusion

This paper has presented a preliminary typologization of how different (meta)morphomic elements can exist together in the same inflectional system or in the paradigm of a single lexeme. This is done in the same vein as recent work (Corbett 2015) dealing with lexical splits more generally. As in that work, all the logically possible configurations (disjoint, overlapping and subset relations) have been found here to exist in natural languages. When cross-classifying morphomes inhabit the same paradigm, this has been found to lead frequently, in the case of the Romance L- and N-morphomes, to the distributional instability of the forms. This is probably due to the multiplication of alternants that this configuration involves and to the increased difficulty, for language users, to find well-defined morphological niches for each of them. A frequent solution, thus, is to reduce the number of alternants (i.e. the orthogonality of the morphomes), either by arranging formerly cross-cutting forms into subset-superset distributions (see Table 23), or by making them disjoint (see Table 8).

In addition, relying on insights from Set Theory, I have typologized what kinds of new morphomic categories may emerge from the cross-classification of morphomes (Union  $[A \cup B]$ , Intersection  $[A \cap B]$ , Relative complement of A in B  $[B-A]$ , Relative complement of B in A  $[A-B]$ , and Symmetric difference  $[A \Delta B]$ ). All of these except the last one have been shown to be able to constitute the domain of analogical change, which suggests that they can, under the right conditions, become morphomes of their

own. The absence of analogical changes that aim at establishing a morphomic category  $A\Delta B$  is an interesting empirical finding that needs further analysis and cross-linguistic validation but which, if solid, would constitute a significant restriction in the area of morphome interactions and more generally in the possible domains for linguistic generalizations and language users' construction of (morphological) categories.

Future research should be aimed at contrasting the validity of this restriction within and beyond Romance and at uncovering other cross-linguistic generalizations concerning morphomes. On this respect, it is usually assumed (e.g. Koonts-Garboden 2016 & Maiden 2018:22) that morphomes must be typologically unique (i.e. that the replication of a pattern in unrelated languages would exclude a morphological pattern from the ranks of morphomes). This seems to me an unnecessary footnote to the definition of the morphome that discourages typological approaches to the phenomenon. It may well be that at a sufficient level of granularity no two morphomes are exactly the same. This would also hold, however, of probably any linguistic category. I believe there is no reason to give up on typological and comparative research in general because of this. Under a sufficiently lax definition, or looking at some narrower aspect, unrelated morphomic structures (e.g. the Koyi Rai morphome in Table 1 and the Romance N-morphome) can indeed be "the same" (e.g. both apply to 3 and SG in the paradigm). Typological work and cross-linguistic generalizations, thus, can and should be attempted if our knowledge of morphomes is to increase beyond its present state.

Another avenue for future research would be to explore whether there are empirical differences in the way morphomes and morphemes interact with other morphological objects. Some recent work has emphasized the gradual nature of (e.g. Smith 2013) or the absence of empirical evidence for (Herce forthcoming) the morpheme/morphome distinction. In this respect, the domain of interactions seems particularly appropriate to try to find differences. Although some morphemes may be so too, morphomes are communicatively largely irrelevant by definition. It would make functional sense if they were free to change their distribution in ways that morphemes cannot do (without jeopardizing information-transfer). For example, orthogonality might well be the ideal configuration of morphemes but may be dispreferred in the case of morphomes. A large cross-linguistic study of paradigmatic analogical changes would be needed, of course, to test this intuition and others.

**Acknowledgements** Thanks are due to Matthew Baerman, Oliver Bond and Ivan Igartua for their comments on early versions of this paper. The financial support of the Basque Government (PRE\_2015\_1\_0175) is also gratefully acknowledged.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge: MIT Press.

- Aronoff, M. (2016). Competition and the lexicon. In A. Elia, C. Iacobini, & M. Voghera (Eds.), *Livelli di Analisi e fenomeni di interfaccia. Atti del XLVII Congresso Internazionale della Società di Linguistica Italiana* (pp. 39–52). Roma: Bulzoni Editore.
- Baerman, M., & Corbett, G. G. (2012). Stem alternations and multiple exponence. *Word Structure*, 5(1), 52–68.
- Bobaljik, J. D., & Sauerland, U. (2018). ABA and the combinatorics of morphological features. *Glossa: A Journal of General Linguistics*, 3.1.
- Luís, A. R. & Bermúdez-Otero, R. (Eds.), 2016. The morphome debate, Oxford: Oxford University Press.
- Barcos, M. A. (2007). *El aragonés ansotano. Estudio lingüístico de Ansó y Fago*. Zaragoza: Gara d'Edizion – Institución „Fernando el Católico”.
- Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.
- Corbett, G. G. (2007). Canonical typology, suppletion, and possible words. *Language*, 83(1), 8–42.
- Corbett, G. G. (2015). Morphosyntactic complexity: a typology of lexical splits. *Language*, 91(1), 145–193.
- Donohue, M. (2004). A grammar of the Skou language of New Guinea. Ms. National University of Singapore.
- Esher, L. (2014). Autonomous morphology and extramorphological coherence. *Morphology*, 24(4), 325–350.
- Haensch, G. (1958). Las hablas de la Alta Ribagorza. *Archivo de filología aragonesa*, 10, 57–194.
- Herce B. (forthcoming). On morphemes and morphomes: exploring the distinction. Word Structure.
- Koonts-Garboden, A. (2016). Thoughts on diagnosing morphomicity: A case study from Ulwa. In A. R. Luís & R. Bermúdez-Otero (Eds.), *The morphome debate* (pp. 228–247). Oxford: Oxford University Press.
- Lahoussois, A. (2009). Koyi Rai: An initial grammatical sketch. *Himalayan linguistics*.
- Lee, J. (1987). Tiwi today. *A study of language change in a contact situation*. Pacific Linguistics, Series C 96.
- Maiden, M. (1992). Irregularity as a determinant of morphological change. *Journal of Linguistics*, 28(2), 285–312.
- Maiden, M. (2004). Morphological autonomy and diachrony. *Yearbook of morphology 2004*, 137–175.
- Maiden, M., Smith, J. C., Goldbach, M., & Hinzlin, M.-O. (Eds.) (2011). *Morphological autonomy: Perspectives from Romance inflectional morphology*. Oxford: Oxford University Press.
- Maiden, M. (2012). A paradox? The morphological history of the Romance present. In S. Gaglia & M.-O. Hinzlin (Eds.), *Inflection and word formation in Romance languages 186*, Amsterdam: John Benjamins.
- Maiden, M. (2018). *The Romance verb: morphomic structure and diachrony*. Oxford: Oxford University Press.
- McCreight, K., & Chvany, C. V. (1991). Geometric representation of paradigms in a modular theory of grammar. *Paradigms: The economy of inflection*, 91–112.
- Montermini, F., & Bonami, O. (2013). Stem spaces and predictability in verbal inflection. *Lingue e Linguaggio*, 12(2), 171–190.
- Nübling, D. (2000). *Prinzipien der Irregularisierung. Eine Kontrastive Analyse von Zehn Verben in Zehn Germanischen Sprachen*. Tübingen: Niemeyer.
- O'Neill, P. (2011). *The Ibero-Romance verb: Allomorphy and the notion of the morphome*. PhD dissertation.
- O'Neill, P. (2012). New perspectives on the effects of yod in Ibero-Romance. *Bulletin of Spanish Studies*, 89(5), 665–697.
- O'Neill, P. (2018). Velar allomorphy in Ibero-Romance: Roots, endings and clashes of morphemes. In M. Bouzouita, I. Sitaridou, & E. Pato (Eds.), *Studies in historical Ibero-Romance morpho-syntax: A descriptive and prescriptive analysis* (pp. 13–46). Amsterdam: John Benjamins.
- Penny, R. J. (2002). *A history of the Spanish language*. Cambridge: Cambridge University Press.
- Pertsova, K. (2011). Grounding systematic syncretism in learning. *Linguistic Inquiry*, 42(2), 225–266.
- Round, E. R. (2013). Rhizomorphemes, meromorphemes, and metamorphemes. In M. Baerman, D. Brown, & G. Corbett (Eds.), *Understanding and measuring morphological complexity* (pp. 29–52). Oxford: Oxford University Press.
- Rupp, J. (1996). *Diccionario chinanteco de San Juan Lealao, Oaxaca*. Tucson: SIL.
- Smith, J. C. (2013). The morphome as a gradient phenomenon: evidence from Romance. In S. Cruschina, M. Maiden, & J. C. Smith (Eds.), *The boundaries of pure morphology* (pp. 247–261). Oxford: Oxford University Press.

- de Sousa, H. (2006). *The Menngwa Dla language of New Guinea*. PhD dissertation, University of Sidney.
- Wheeler, M. W. (2011). The evolution of a morpheme in Catalan verb inflection. In M. Maiden et al. (Eds.), *Morphological autonomy: perspectives from Romance inflectional morphology* (pp. 183–209).
- Wilbur, J. (2015). *A grammar of Pite Saami*. Berlin: Language Science Press.
- Wunderlich, D. (2012). Polarity and constraints in paradigmatic distinctness. In J. Trommer (Ed.), *The morphology and phonology of exponence* (pp. 160–195). Oxford: Oxford University Press.
- Zipf, G. K. (1935). *The psycho-biology of language*.