13
Conditioned epenthesis in Romance

Mark Aronoff and Lori Repetti

13.1 Introduction

Most modern discussion of epenthesis has concentrated on types that are phonetically or phonologically motivated. We show that the factors governing individual phenomena may go beyond phonetics and phonology to morphology and morphosyntax. We first discuss cases from Romance languages where a language has more than one epenthetic segment. Here, although the motivation for epenthesis may be phonological, the choice of one segment over another is determined by further factors: it can be morphologically conditioned or influenced by morphosyntax. Further up the scale, the insertion of a meaningless syllable—including stem extenders in Formentera Catalan, Italian <*isc*>, and Spanish antesuffixal interfixes—may have little if any phonological motivation, but is influenced by morphology. We conclude that the notion of epenthesis should be broadened from the purely phonological to include morphological and morphosyntactic conditions. Many of the phenomena that we discuss have been accounted for in terms of allomorphy, with allomorphs that are lexically listed. ~~Furthermore~~, an analysis based on allomorphy is always descriptive and never explanatory. The purely allomorphic approach gives up on the possibility of finding more widely applicable constraints on the form and presence of the epenthetic segment, and hence misses a broad generalization: when more than one epenthetic segment is possible, the choice is predictable within specific morphological or morphosyntactic contexts.[1]

Epenthesis is 'the interposition of a letter or syllable in the midst of a word' (Smith 1656, the earliest citation in *OED*). The term and concept are ancient, dating to Classical Greek rhetoric. Although most modern treatments have concentrated on cases of epenthesis that are phonologically motivated, we show in this chapter

[1] More than anyone else, Martin Maiden has revealed the power of the morphological explanation of complex phenomena throughout the Romance languages, most notably in his monumental volume, *The Romance Verb* (Maiden 2018a). We are delighted to honour him with this short contribution, whose focus is a morphological explanation of a phenomenon that has usually been viewed through a phonological lens.

Mark Aronoff and Lori Repetti, *Conditioned epenthesis in Romance*. In: *Periphrasis and Inflexion in Diachrony*. Edited by Adam Ledgeway, John Charles Smith, and Nigel Vincent, Oxford University Press.
 DOI: 10.1093/oso/9780198870807.003.0013

that the factors governing it range from the purely articulatory to those that lie beyond phonology: morphology, morphosyntax, and as far as the lexicon. Here we exclude lexical conditions, for the sake of brevity.

The insertion of an 'unetymological' segment is rampant throughout the world's languages and goes under a wondrous plethora of ancient names, from *epenthesis* (AGk. *ἐπένθεσις* < *ἐπί* 'in addition' + *ἐν* 'in' + *θέσις* 'placing') to *anaptyxis* (also AGk. *ἀνάπτυξις* < *ἀνά* 'up' + *ptyxis* 'folding' 'unfolding', reserved for vowels) to *svarabhakti* (Sanskrit *svara* 'vowel' *bhakti* 'separation' 'vowel separation') and its consonantal counterpart *vyanjanabhakti*. In the phonetic tradition, there are less ancient terms for the phenomenon: *transitional segment*, especially *transitional vowel*, and most recently, two new terms, *excrescent vowel* (Levin 1987) and *intrusion vowel* (Hall 2006). In this contribution, we make the case for a unified approach to this cornucopia, while concentrating our attention on a limited set of data from Romance languages that lie on the morphological part of the spectrum.

The conditions under which all these insertions occur run the gamut from purely phonetic to morphosyntactic. What unifies the phenomena is the fact that something is inserted between segments. What makes for the diversity are the sorts of conditions under which the insertion takes place. Our innovative contribution is the observation that both the type of condition and the nature of what is inserted are tied to what we used to call the *linguistic level* of the insertion.

13.1.1 Intrusive vowels

One of the more remarkable types of insertion is found at the articulatory level and is attributed to the articulatory intrusion between consonants. Hence the term *intrusive*, as with the intrusive stops in English words like *sense* ([sɛns] or [sɛnts]) and *temse* ([tɛms] or [tɛmps]) 'sieve'. Martin Maiden (1995:242) notes the same insertion of [t] in sequences of sonorant + [s] in various Romance varieties spoken in central and southern Italy, e.g., Umbrian [pɛ́ntso] < *pɛnso 'I think', [fáltso] < *falsu 'false'. The fact that these intrusions are specific to certain varieties shows that they are not automatic, but part of the linguistic system. Hall (2006) shows in some detail that intrusive vowels (as in the pronunciation of *arm* as [árəm] in some varieties of English) often do not rise to the status of segments: they do not count as syllabic nuclei in the phonological timing calculus. 'Ordinary epenthetic vowels, however, are syllable nuclei' (Hall 2006:388). Bertil Malmberg (1950) was the first to document intrusive vowels (in Argentinian Spanish), and they have since been observed in other varieties of Spanish and in French (e.g., Colantoni and Steele 2004). Miatto (2020) shows that certain word-final inserted vowels in Italian also qualify as intrusive in this sense, not having segmental status in the phonology.

Hall emphasizes that the main difference between intrusive and canonical epenthetic vowels is that the latter count as full segments. While both types can

have a single value throughout the language, so-called copy vowels, where the value of the vowel is taken from that of a nearby full vowel, may be more common in cases of intrusion. This is in line with the instability of excrescent vowels observed by Levin (1987). They are often optional or correlated with the speed of speech (less likely in faster speech). Researchers find that native speakers are also often unaware of the presence of intrusive vowels (Colantoni and Steele 2004).

13.1.2 Phonological epenthesis

Phonological epenthesis, by contrast with intrusion, is the insertion of meaningless phonological material whose appearance is motivated by phonology (to repair an illegal structure), and whose quality is usually unmarked in the language (Kitto and de Lacy 1999). Consider the example in (1), where phonologically impermissible word-initial *s-stop* clusters are repaired by epenthesizing an initial vowel [e], which is unmarked in the language.[2] (Throughout this chapter, the epenthetic segment is underlined.)

(1) Spanish: <stop> [e̲stóp]

As with intrusion, the quality of the epenthetic vowel may vary depending on a number of phonological and prosodic factors. Factors include its position, e.g., English loanwords in Bengali (Broselow 2015): *school* > [i̲skul], *glass* > [ge̲láʃ]; the surrounding phonetic environment, e.g., Afrikaans loanwords in Sotho (Rose and Demuth 2006):/blɪk/> [bʊ̲léke] 'tin can', /truwn/> [tu̲róni] 'throne'; and the neighbouring vowels in 'copy epenthesis', e.g., English borrowings in American varieties of Italian (Repetti 2012): *washtub* [veʃʃetúbbu], *cocktail* [kokkɔ̲télla], *popcorn* [pappa̲kɔ́rno], among others. In all of these cases, the choice of the inserted segment is determined phonologically.

In a number of cases, however, the quality of the inserted segment does not seem to be determined phonologically (Moradi 2017). Put another way, the presence of a vowel is motivated phonologically, but the exact quality of the epenthesized vowel is determined by other factors. For example, in the Romance variety of San Marino (Michelotti 2008), we find two possible epenthetic vowels to satisfy restrictions on word-final clusters: [ɪ]/[ɐ].

(2) San Marino
a. /ojm/ > [ójmɪ̲] 'elm'
b. /dɔrm/ > [dɔ́:rmɐ̲] 'sleep.PRS.IND.3SG/PL' ~~(Michelotti 2008)~~

[2] Archangeli (1988) and others argue that /e/ is the maximally underspecified vowel in Spanish, and others show that /e/ is the most frequent phoneme in Spanish (Guirao and García Jurado 1990). Thanks to a reviewer for pointing this out.

In San Marino, [ɪ] is the default epenthetic vowel (2a),[3] but [ɐ] occurs in word-final position with certain verbs ending in unacceptable clusters (2b). The choice between [ɪ] and [ɐ] is made on the basis of morphosyntactic considerations: [ɪ] is the usual epenthetic vowel, except in certain morphosyntactic contexts when epenthetic [ɐ] in employed instead (see §13.2.2 for more details). We see from this example that the factors conditioning epenthesis are richer than previously noted and go beyond phonological factors to morphological and morphosyntactic factors. In what follows, we present a number of examples from Romance languages where the motivation for epenthesis is phonological, but the choice of segment is determined by these factors.

This chapter is organized as follows. In §13.2 we introduce five case studies of morphologically conditioned epenthesis from Romance languages. We provide a review of previous accounts of the data in §13.3, including allomorphic accounts (§13.3.1) and epenthesis accounts (§13.3.2). We present our analysis in §13.3.3, and then discuss other types of insertion (§13.4). We summarize our proposal and conclude in §13.5.

13.2 Case studies in Romance languages

In this section we present five case studies from Romance languages in which two phonologically distinct epenthetic segments may be used to satisfy phonological constraints. Crucially, the choice between the two segments is influenced by morphology or morphosyntax rather than phonology.

13.2.1 Brazilian Portuguese

Brazilian Portuguese chooses between two epenthetic segments in resolving vowel hiatus: [j] is the default (3a), except when hiatus results from the addition of an affix, where [z] is used instead at the morphological juncture (3b) (Bachrach and Wagner 2007; Garcia 2017). [i] is the default epenthetic vowel in Brazilian Portuguese: it is used to break up impermissible morpheme-internal clusters (*psicologia* 'psychology' can be pronounced [p̲isikoloʒíɐ]), and some cases of consonant-final loanwords are adapted with insertion of final [i]: Eng. *blog* > [blɔ́gi] (Xavier 2013 reported in Artes 2016:114). The default epenthetic consonant [j] is the non-vocalic counterpart of [i]. At some level, we can think of them as the same element.

(3) Brazilian Portuguese
   a. *Correa* [koréja] '(name)'
   b. /sofá + iɲu/ > [sofaz̲íɲu] 'sofa (DIM)'
      /kafé + al/ > [kafez̲ál] 'coffee grove'[4]

[3] The quality of the vowel is either [ɪ] or [i] which occurs in free variation (Michelotti 2008:332).
[4] The /al/ suffixes following the name of a plant means the place where the plant is cultivated.

Epenthetic [z] appears instead of the expected [j] in certain hiatus contexts between morphemes, e.g., BrPt./kafé + ĩɲu/ 'coffee' + DIM > [kafezíɲu], **[kafeíɲu] 'little coffee (espresso)', and only in hiatus contexts, e.g., BrPt. /zebr-a + iɲa/'zebra' + DIM > [zebríɲa], [zebrazĩɲa], **[zebrzíɲa] 'little zebra'. Bachrach and Wagner (2007:7) show that the use of forms without epenthesis ([zebríɲa]) or with epenthesis ([zebrazíɲa]) is not only influenced by phonology, but also reflects the syntactic attachment site of the suffix. They provide details about the morphology, syntax, and semantics of the diminutive and other suffixes. Thinking more within the tradition of generative phonology, we may portray the difference between the members of pairs like [zebríɲa] and [zebrazíɲa] as differences in lexical levels (I and II in Lexical Phonology) or boundaries (+ and #) in SPE terms. Going further back, they are analogous to Sapir and Bloomfield's two English *-er* suffixes, where one attaches as if to a word ([síŋəɹ] vs [lɔ́ŋgəɹ]) (Sapir [1925] 1949:43).[5] The most important point for the moment is that the epenthetic consonant between two morphemes in hiatus, [z], is not the default purely phonologically conditioned segment in the same hiatus context ([j]). The addition of the morphological condition changes the environment. We will return to the Brazilian Portuguese case at the end of §13.3, where we discuss additional morphological complexities.

13.2.2 San Marino

The Romance language spoken in the independent republic of San Marino in northern Italy chooses between two epenthetic vowels in word-final position, [ɪ] and [ɐ] (see note 3), as illustrated in (2). Both epenthetic vowels are used with unsyllabifiable consonant clusters word-initially that result from syncope of pretonic front vowels (Michelotti 2008:126): [i] is inserted within a consonant–sonorant–consonant cluster /krsu:/ > [kr̲isú:] 'grown'; [ɐ] is used in initial position before a sonorant–consonant sequence /lgɛ:/ > [ɐ̲lgé:] 'to tie'. Word-finally, [ɪ] is inserted with nouns (/lɛ:dr/ > [lé:dr̲ɪ] 'thief'), adjectives (/sɐlba:dg/ > [sɐlbá:dg̲ɪ] 'wild'), and adverbs (/sɛ̲impr/ > [sɛ́i̲mpr̲ɪ] 'always'). The word-final situation with verbs is more complex and depends on the person, number, and conjugation class of the verb form. Verbs normally have no suffix marking the first-/second-person singular of the present indicative, regardless of their conjugation class (4).[6]

(4) a. 1st conjugation class: [ba:l] 'dance.PRS.IND.1/2SG'
b. 4th conjugation class: [fnɪs] 'finish.PRS.IND.1/2SG'

[5] Sapir attributes the analysis to a personal communication from Bloomfield. The distinction does not hold in many regional varieties.

[6] Michelotti (2008:331) categorizes verbs into four conjugation classes. For the phenomena discussed here, the second, third, and fourth conjugation classes pattern the same way, and we will use data from the fourth conjugation class to illustrate the patterns of the three classes. In addition, we are not including subject clitic pronouns which are obligatory in some contexts.

However, if the verb ends in an illegal cluster, epenthetic [ɪ] is inserted at the end of the verb (5).

(5) a. 1st conjugation class: [úːrlɪ̲] 'scream.PRS.IND.1/2SG'
b. 4th conjugation class: [íːrvɪ̲] 'open.PRS.IND.1/2SG'

For third-person singular and plural present indicative verbs only, the first conjugation class differs from all others. The theme vowel [ɐ] marks the third-person singular and plural of first conjugation verbs, while the third-person singular and plural of the other conjugation class verbs have a different theme vowel and no suffix (6).

(6) a. 1st conjugation class: [báːlɐ] 'dance.PRS.IND.3SG/PL'
[úːrlɐ] 'scream.PRS.IND.3SG/PL'
b. 4th conjugation class: [fnɪs] 'finish.PRS.IND.3SG/PL'

However, if the verb ends in an illegal cluster, epenthetic [ɐ] rather than the expected [ɪ] is inserted at the end of the verb in the other conjugation classes, as in (7).

(7) 4th conjugation class: [íːrvɐ̲] (**[íːrvɪ̲]) 'open.PRS.IND.3SG/PL'

A summary of the relevant verb forms is provided in Table 13.1. The ɐ-bearing cells are shaded.

Table 13.1 San Marino verbs

		acceptable ending	illicit final cluster repair
1SG/2SG	1st conj. class	/bal/ [baːl]	/url/ [úːrlɪ]
	4th conj. class	/fnɪs/ [fnɪs]	/irv/ [íːrvɪ̲]
3SG/3PL	1st conj. class	/bal + ɐ/ [báːlɐ]	/url + ɐ/ [úːrlɐ]
	4th conj. class	/fnɪs/ [fnɪs]	/irv/ [íːrvɐ̲]

We analyse the [ɪ] of both [úːrlɪ̲] 'scream.PRS.IND.1/2SG' (first conjugation class) and [íːrvɪ̲] 'open.PRS.IND.1/2SG' (fourth conjugation class) as epenthetic. Furthermore, the [ɐ] of [úːrlɐ] 'scream.PRS.IND.3SG/PL' (first conjugation class) is an inflexional morpheme (cf. [báːlɐ] 'dance.PRS.IND.3SG/PL'), while the [ɐ] of [íːrvɐ̲] 'open.PRS.IND.3SG/PL' (fourth conjugation class) is an epenthetic vowel.

In San Marino, [ɪ] is the default epenthetic vowel, but [ɐ] is used to satisfy phonological constraints on word-final position in specific morphosyntactic contexts, namely with second-, third-, and fourth-conjugation third-person singular and plural verbs ending in an unacceptable cluster. The data in (6) clearly show that [ɐ] does not mark third-person singular and plural in general, but only for first-conjugation verbs. However, with non-first-conjugation verbs, if an epenthetic

vowel is needed, [ɐ] is used. Its presence is phonologically motivated, but its quality is not. The choice between canonical [ɪ] and non-canonical [ɐ] is made on the basis of morphosyntactic considerations: [ɪ] is the usual epenthetic vowel, except with third-person singular and plural verbs, which instead employ epenthetic [ɐ] in final position.

The next examples similarly illustrate cases in which the choice between two epenthetic vowels is made based on morphosyntactic considerations.

13.2.3 Paduan

Paduan, a Romance variety spoken in the province of Padua in the Veneto region of northern Italy, also chooses between two epenthetic segments to satisfy syllable constraints: [e] is the default vowel used in most cases (8a–b), except at the right edge of a phrase consisting of a verb + pronoun, when [o] is used (8c–d) (Cardinaletti and Repetti 2007; 2008). The choice of epenthetic vowel ([e] or [o]) is determined by its position within a morphosyntactic context.

(8) Paduan
- a. /l maɲa/ > [e̲l máɲa] 'he eats'
- b. /t maɲi/ > [te̲ máɲi] 'you eat'
- c. /maɲa l/ > [máɲe lo̲] 'eats he' > 'does he eat?'
- d. /maɲi t/ > [máɲi to̲] 'eat you' > 'do you eat?'

In Paduan, verb+clitic and clitic+verb structures form a phonologicalpPhrase (see Selkirk 1995 and Peperkamp 1997 for the prosodization of clitics) and are subject to certain constraints (Cardinaletti and Repetti 2008). In (9a–b) the default epenthetic vowel [e] is used to syllabify the proclitic. Its position before the /l/ in (8a) is determined by an alignment constraint that accounts for the 'peripherality of epenthesis' whereby a host and its clitic must be adjacent (Bonet and Lloret 2005). However, in (8b) its position after the /t/ is determined instead by (more general) constraints on coda consonants: /t/ is never an acceptable syllable coda in Paduan: **[e̲t máɲi]. Crucially, epenthetic [e] is used in both contexts. In (8c–d), however, the special epenthetic vowel [o] is used at the end of a verb + enclitic pronoun unit.[7] The default epenthetic vowel [e] is avoided in these contexts (**[máɲe le̲], **[máɲi te̲]), and a special morphosynctactically restricted epenthetic vowel [o] is used. Its presence satisfies phonological constraints, but its quality is not the usual one. Importantly, we are not claiming that the final [o] of [máɲe lo̲] and [máɲi to̲] is a morpheme; we are arguing that it is an epenthetic vowel selected for a special morphologically circumscribed position. As with the data from San Marino, a special epenthetic vowel is used in a restricted morphosyntactic position.

[7] Paduan does not allow final /t/ or /l/, only final nasals (Zamboni 1981:34).

13.2.4 Italian

Italian uses two epenthetic segments to repair consonant cluster violations, and the choice between them is determined by morphosyntactic considerations: [i] is the default vowel used in most cases (9a), including with the masculine singular definite article (9b). However, at the right edge of the masculine singular definite article, epenthetic [o] is used instead of [i] (9c) (Cardinaletti and Repetti 2008; Repetti 2012; 2021).

(9) Italian
- a. spoken varieties: /psikɔ́logo/ > [p̲isikɔlogo] 'psychologist'
  historical change: *alisna* > [lés̲ina] 'awl'
  fixed spoken phrases: /per skritto/ > [per i̲skrítto] 'written'
  American varieties of Italian: *box* [bók̲isa] (Di Vita 1931)
- b. definite article: /l kane/ > [i̲l káne] 'the dog'
- c. definite article: /l spɛkkjo/ > [lo̲ spɛ́kkjo] 'the mirror'

This analysis of the Italian masculine singular definite article rests on the assumption that the input form of the definite article is /l/, and the distribution of the three surface forms—[l], [il], [lo]—is predictable (for more details, see Repetti 2021).

(10) Italian masculine singular definite article
- a. [l] before a vowel: *l'amico* 'the friend'
- b. [il] before a single consonant (except those in (c)) and certain clusters (Cl, Cr, Cj, Cw): *il bambino* 'the child'
- c. [lo] before /t͡s/, /d͡z/, /ʃ/, /ɲ/, /ʎ/ and sC clusters: *lo specchio* 'the mirror'

If we posit an input form /l/, we can easily account for the form in (10a) since no changes to the input form are necessary: the /l/ can be syllabified as the onset of the syllable: /l amiko/ > [la.mí.ko] 'the friend'. An epenthetic vowel is necessary in (10b) since an onset cluster cannot begin with /l/: /l bambino/ > [i̲l.bam.bí.no], and the default epenthetic vowel is used. The position of the epenthetic vowel is accounted for by means of the alignment constraint discussed above, resulting in peripheral epenthesis. In (10c), an epenthetic vowel is needed, and it is placed after the /l/, in violation of the alignment constraint, because of more highly ranked constraints on syllable structure:[8] /l spɛkkjo / > [lo̲s.pɛ́k.kjo] 'the mirror'. However, the default vowel [i] is not used. Instead, we find [o]. As in San Marino and Paduan, the motivation for vowel epenthesis is phonological, but the quality of the

[8] The same analysis holds for ~~sC clusters and~~ the 'inherently long' consonants in (10c):/l t͡sμio/> [lo̲t t͡sí.o] 'the uncle'. See Repetti (2021) for details.

epenthetic segment is not the usual one.[9] Of course, the choice of the vowel [o] is no accident. The default masculine singular ending is this same vowel. The close parallel with San Marino is telling. In that language, the default third-person singular present verbal ending [ɐ] overrides the default epenthetic vowel [ɪ] under morphologically circumscribed conditions. Here, the vowel of the default masculine singular nominal ending *-o* overrides the default epenthetic vowel [e] (under different morphologically circumscribed conditions). In both cases, a morphologically specific epenthetic vowel that is phonologically identical to a more general morpheme overrides a more general phonological default.

13.2.5 Catalan varieties

Alguerès, the variety of Catalan spoken in Sardinia, uses two epenthetic vowels to satisfy syllable constraints: [i] is the default value used between words (11a), and [u] is used before the masculine plural suffix /s/ if an epenthetic vowel is needed (11b) (Loporcaro 1997c). Note that if an epenthetic vowel is not needed, the [u] does not appear before the plural marker: /mɔlt + s/ > [mɔlts] 'dead.MPL'.

(11) Alguerès
  a. /amik meu/ lit. 'friend my' > [amík i̲ méu] 'my friend'
  b. /fresk + s/ 'cool.M + PL' > [fréskus̲] 'cool (MPL)'

Another variety of Catalan, Pallarès, spoken in Pallars in Catalonia, similarly chooses between two epenthetic vowels: [e] and [o] (Artes 2016). The former is the default vowel used to satisfy syllable constraints in most cases (12a). Its quality in central Catalan is [ə], but in Pallarès it is usually [e] (Artes 2016:148f.) (see Jiménez 2008 and Lloret and Jiménez 2008 for more on epenthetic vowel quality in Catalan varieties). The special epenthetic vowel [o] ([u] in central Catalan) is used before the masculine plural suffix /s/ (namely, with sibilant-final nouns and adjectives) (12b). As in Alguerès, Pallarès epenthetic [o] is not used with the masculine plural suffix if an epenthetic vowel is not needed: /gat + s/ > [gats] (**[gatos̲]) 'cats'.

(12) Pallarès Catalan
  a. [e̲]Spielberg 'Spielberg (film director)'
  b. /gos + s/ > [gósos̲] 'dogs'

In the case studies above, we have seen that the choice between two epenthetic segments is made based on morphosyntactic considerations (morpheme boundaries §13.2.1, verb classes §13.2.2, phrases involving clitic pronouns §13.2.3) and

[9] Note that there is a difference between the Italian (9)–(10) and Paduan (8) patterns. In Italian, the right edge of the definite article requires the selection of the special epenthetic vowel: /l/ > [lo̲] (10c), even though ~~that~~ is not at the right edge of the phrase, while in Paduan, the right edge of the subject pronoun only triggers the selection of the special epenthetic vowel when it is at the end of the phrase: /l/> [lo̲] (8c) and /t/> [to̲] (8d), but not phrase-internally: /l/> [e̲l] (8a) and /t/> [te̲] (8b).

on specific morphemes (definite article §13.2.4, plural suffix §13.2.5). In these cases, the default epenthetic vowel is not used, but instead a special segment is selected. In the next section we discuss possible analyses of these phenomena.

13.3 Accounts of the phenomena

Various analyses of the phenomena described above have been proposed in the literature. We consider two approaches: allomorphy (§13.3.1) and epenthesis (§13.3.2), and we conclude that the inserted segment is indeed epenthetic, and its quality can be morphologically conditioned (§13.3.3).[10]

13.3.1 Allomorphic solutions

Many of the phenomena outlined in §13.2 have been addressed in the literature as cases of allomorphy. For example, the San Marino first- and second-person singular verbs end either in Ø or/ ɪ/ (Table 13.1), and Michelotti (2008:332) refers to /ɪ/ as 'a phonologically conditioned allomorph of the desinence occurring in complementary distribution with -Ø'. Similarly, third-person singular and plural verbs of the fourth conjugation class end either in Ø or/ɐ/, and Michelotti (2008:335) refers to /ɐ/ as 'a complementary allomorph of the desinence -Ø'.[11] The different Paduan proclitic vs enclitic forms (8) have also been attributed to allomorphy: some argue that there is a proclitic paradigm and an enclitic paradigm of pronouns (Munaro 1999) or an 'interrogative conjugation' (Zamboni 1974), while others claim the proclitic is a pronoun, and the enclitic is an inflexional suffix (Zamboni 1974; Benincà and Vanelli 1982; Benincà 1983; Poletto 2000). The Italian case (10) has also been described in terms of allomorphy. Many scholars have accounted for the distribution of the three forms of the masculine singular definite article by positing various listed forms: /l, il, lo/ (Dressler 1985; McCrary 2004; McCrary Kambourakis 2007; Garrapa 2009; 2011), /il, lo/ (Davis 1990; Marotta 1993; Morelli 1999; Krämer 2009), /l, il/ (Vanelli 1992; Mascaró 1996; Tranel and Del Gobbo 2002). The appropriate allomorph is selected based on the phonological

[10] Many authors propose historical accounts of the synchronic facts, including Michelotti (2008) for San Marino (§13.2.2), Vanelli (1984; 1987) for Paduan (§13.2.3), Gröber (1877), Ambrosini (1978), Vanelli (1992), Renzi (1993) for Italian (§13.2.4), Loporcaro (1997c) for Alguerès (§13.2.5), and Artes (2016) for Pallarès (§13.2.5). A discussion of these approaches is beyond the scope of this chapter.

[11] Another possible account of the San Marino data is that this is an example of heteroclisis whereby the inflexional paradigm of a particular lexeme involves more than one inflexional class (Stump 2006). In the case of San Marino, this would mean that some second-, third-, and fourth-conjugation verbs adopt the inflexional suffix of first-conjugation verbs for the third-person singular and plural of the present indicative. However, this approach does not appear to apply because the unexpected final vowel only occurs in a particular phonological context, namely with an illicit final cluster. Thanks to a reviewer for pointing out this option to us.

context. As for the Catalan data (§13.2.5), Bonet, Lloret, and Mascaró (2007) have proposed allomorphs of the masculine noun suffixes ('gender allomorphy'). The listed and ranked lexical entries for masculine gender include /Ø > u/.[12] The first allomorph is the one used, unless the plural suffix /s/ cannot be adjacent to the noun stem for phonological reasons, in which case the second is selected.

Phonologically conditioned allomorphy is the most common account of the phenomena described in §13.2 and can be invoked to describe each case. However, this approach suffers from a serious problem: it fails to take into account the fact that the restricted 'allomorph' falls in with a larger morphological pattern in the language. The fact that [ʊ] is used in the restricted context in San Marino, [o] in Paduan and Italian, and [u]/[o] in Catalan is predictable, but simply listing all the allomorphs from which to choose results in an unnecessarily complex grammar. We expand on this in §13.3.2.

We are not arguing that the listing of allomorphs does not ever exist. Allomorphs whose forms are unpredictable from other aspects of the language must clearly be listed in the lexicon, as with the English indefinite article *a/an*, definite article [ðə]/[ði], and suppletive forms of the verb BE. But when the restricted 'allomorph' falls in with a larger pattern in the language, lexical listing prevents us from seeing that pattern. In general methodological terms, lexically listed allomorphy is a brute-force tool for describing the language. Listing allomorphs can never have any explanatory value. Positing a restricted allomorph should always be the solution of last resort, to be called upon only when all else fails.

13.3.2 Epenthesis

The phenomena outlined in §13.2 have also been accounted for as instances of epenthesis. In the case of Brazilian Portuguese (§13.2.1), Bachrach and Wagner (2007:8), Garcia (2017:47), and others argue for a [z]-insertion process when the addition of an affix results in hiatus,[13] but they do not address its quality. In other cases of hiatus resolution, [j] is used.

The Catalan facts (§13.2.5) have also been accounted for as cases of epenthesis. Alguerès uses two epenthetic segments which Loporcaro (1997c) refers to as 'i-epenthesis' vs 'u-epenthesis'. The quality of epenthetic [i] is only recently stabilized from a more variable mid-central vowel quality reported less than a century ago (Loporcaro 1997c:215–217). On the other hand, he accounts for the quality of [u] as a case of 'reanalysis leading to rule inversion', i.e., historical [u]-deletion was blocked in contexts where deletion would have resulted in a disallowed cluster.

[12] Bonet, Lloret, and Mascaró (2007) espouse a framework in which allomorphs are not simply listed, but ranked. The ranking does not follow simple elsewhere principles but must be specified.

[13] Bachrach and Wagner (2007:7) further argue that the forms with or without the /z/ differ in the point of attachment.

At that point it was reanalysed as [u]-insertion (Loporcaro 1997c:217), Mascaró (1985a; 1985b), Lloret and Viaplana (1992), Wheeler (2005), Bonet et al. (2007), Artes (2016), and others have also proposed that Catalan [u]/[o] is epenthetic~~: (in another approach, Zimmermann (2019) refers to Catalan [u] as a 'ghost' whose appearance is lexically determined)~~, but they argue that its quality is morphologically conditioned. Various proposals regarding the role of morphology in determining the quality of this special epenthetic segment have been offered. Artes's (2016) detailed proposal is that the [o] in (12b) is an inflexional morpheme whose floating features of the input are parsed only when phonotactic requirements make that necessary (p. 269).

The Italian definite article [lo] (§13.2.4) has been analysed as the result of epenthesis by Vanelli (1992) and Tranel and Del Gobbo (2002) who conclude that the quality of the vowel is determined morphologically, and by Cardinaletti and Repetti (2007; 2008), Repetti (2012), and Repetti (2021) who come to the same conclusion for both the Paduan data (§13.2.3) and the Italian definite article (§13.2.4). The latter authors invoke the privileged position at the end of a word or phrase where inflexional information is located to account for the Paduan and Italian facts. They start with the premise that a vowel representing a morphologically marked category such as PL is avoided as an epenthetic vowel in final position. The vowels /e/ and /i/ mark plural number in Paduan and Italian nominals and so are avoided as epenthetic vowels in final position, where inflexional markers are usually found. For the Italian definite article, they posit the underlying form of the article as /l/, and the surface forms ([l], [il], and [lo]) derive from /l/ by means of independently attested processes. The fact that [o] is used at the end of the article (rather than usual epenthetic vowel [i]) is justified on the grounds that [o] is the morphologically neutral final vowel in Italian nominals (Ferrari 2005), and non-inflexional [o] is found at the end of other morphemes in the nominal domain: indefinite and negative pronouns and adjectives, as well as adverbs (*uno* 'one', *qualcuno* 'someone', *altro* 'other', *nessuno* 'no one', *ciascuno* 'each one', *tutto* 'all', *tanto* 'so much', *poco* 'little', *molto* 'much') and the predicate clitic pronoun *lo*, as in *Maria è simpatica, e anche Giovanna lo è* 'Maria is nice, and Giovanna is [lit. "is it"] too' (Cardinaletti and Repetti 2007). Similarly, in Paduan a morphologically marked vowel in final position ([e] which marks feminine plural nouns) is avoided, and instead a different vowel is used: [o].

In the next section we build on these analyses.

13.3.3 Interim conclusion: morphologically conditioned epenthetic segment quality

Our interim conclusion is that morphology and morphosyntax play a role in selecting the epenthetic segment used in a particular position; however, that segment is an epenthetic segment and not a morph.

For San Marino, [ɪ] is the default epenthetic segment, but [ɐ] is used instead in a particular context: at the end of the third-person singular and plural verb because, we contend, [ɐ] is the most frequent vowel found in that context. Michellotti (2008:335) argues for San Marino that the [ɐ] in [i:rvɐ̲] 'open.PRS.IND.3SG/PL' is an allomorph of Ø (as in [fnɪs] 'finish.PRS.IND.3SG/PL') that is used after a final unsyllabifiable consonant cluster. He justifies the use of this particular segment by saying that it is used by analogy with the third-person singular/plural suffix /ɐ/ of the first conjugation class. However, this approach misses a more general point: the restricted 'allomorph'—[ɐ]—happens to be the most frequent vowel used in final position with all third-person singular/plural verbs in all conjugation classes and tense/aspect/mood specifications: [bá:lɐ] 'dance.PRS.IND.3SG/PL' (first conjugation class), [durmí:vɐ] 'sleep.IPFV.IND.3SG/PL (all conjugation classes), [mɔ́:rɐ] 'die.PRS.SBJV.3SG/PL' (all conjugation classes), [sɐrí:ɐ] 'be.COND.3SG/PL' (all conjugation classes) ~~(~~[ɐ] is not the final segment in the third-person singular/plural imperfect subjunctive, perfect indicative, and future forms~~)~~.

With Paduan enclitics and the Italian definite article, [o] is used instead of the default epenthetic vowel ([e] and [i], respectively), because [o] is the morphologically neutral vowel in final position outside the verbal domain, as Ferrari (2005) illustrates for Italian, and Zamboni (1988) for northern Italian varieties. Zamboni (1988) reports a pattern of restitution of historically deleted word-final vowels in north-eastern Italian dialects: the 'restored' vowel is [o], the 'morphological free and neutral vowel' (Zamboni 1988:254), and it is found not only at the end of masculine nouns and adjectives, but also feminine nouns, third-person singular verbs, adverbs, and particles.[14] Two clarifications are in order. First, we repeat that we are not claiming that the final [o] of the Paduan enclitics and the Italian definite article is a morpheme; we are arguing that it is an epenthetic vowel selected for a special position. As with the data from San Marino, a special epenthetic vowel is reserved for phrase-final position, a morphosyntactically salient position. Furthermore, we are not claiming that all cases of final [o] are epenthetic: in most cases, the final [o] of a noun or adjective represents a morpheme and is not epenthetic.

We analyse the Catalan facts in a similar way: the [o]/[u] found in some cases with the plural suffix is the most neutral vowel available for stem-final position in nouns, while the default epenthetic segment [e]/[ə] is a marked vowel in nominals. In Pallarès, [e] is the second most common vowel marking feminine nouns (after [a]) and is used before plural [s] with feminine nouns: *cas*[a]/*cas*[e]*s* 'house/houses' (Artes 2016:118). In other varieties of Catalan, [ə] marks feminine nominals.

To sum up our proposal: if an epenthetic segment is needed for phonological reasons in a morphologically salient position, the usual epenthetic segment might not be used, and a special vowel can be employed instead. In San Marino, the special epenthetic vowel used in final position with third-person singular/plural verbs

[14] See Tekavčić (1977:459–467) for morphosyntactic conditioning of [o] restoration after apocope.

is [ɐ] because that is the most frequent vowel in that position, and in Paduan, Italian, and Catalan, the special epenthetic vowel used in final position with nominals is [o]/[u] because [o]/[u] does not represent a morphologically marked category in final position in nominals.

In each of these cases, we agree with Artes (2016:266f.) that 'the use of regular epenthesis is discarded in word-final position (the location for inflection) to avoid mismatches between phonological and morphological structure'. We propose that the quality of the restricted epenthetic segment is the same as the morphologically neutral morph or the most frequent one in that particular context, but the inserted segment is not a morph.

The one case that we have not discussed is Brazilian Portuguese /z/. Here we have no parallels to fall back on: /z/ does not seem to enjoy any special morphological status outside hiatus contexts (though see the discussion of Spanish antesuffixes in §13.4.3 for a possible etymology). According to Bachrach and Wagner (2007:8), /z/ is epenthesized 'whenever adding an affix creates a hiatus' and is thus morphologically conditioned. They note that an allomorphic analysis 'would require postulating two allomorphs for all affixes', showing that /z/ is indeed epenthetic.

The complexity of the Brazilian Portuguese phenomenon lies elsewhere. First, if a noun lacks a theme vowel, the [z] is inserted even though it does not fall between vowels and is therefore not in a true phonological hiatus context: BrPt. *flor* 'flower', *florzinha* 'flowerlet'. Bachrach and Wagner posit an abstract theme vowel here, whose sole justification is to save the hiatus analysis. Notably, though, only sibilant and rhotic consonants are allowed word-finally in the language. A ~~brute-facts~~ analysis would extend the 'hiatus' context to these consonants as well or might morphologize it: [z] is inserted between a noun stem and a Bloomfield–Sapirian word-level vowel-initial suffix.

Evidence for this morphologized analysis is found in an intriguing observation that lies at the heart of Bachrach and Wagner's article, revealed in pairs like BrPt. *zebrinha* and *zebrazinha*, both 'little zebra', though slightly different in connotation and distribution: the word *zebra* in Brazilian Portuguese can denote an unexpected outcome (most often of a football game). An informal survey of internet citations reveals that *zebrinha* is more likely to denote a small zebra and is often found in child-directed contexts, while *zebrazinha* is more likely to denote an unexpected outcome. As noted above, in cases like *zebrazinha*, in Bloomfield–Sapirian terms, the suffix appears to be attached to a whole word, as shown by the retention of the theme vowel *-a*, while in *zebrinha* the affix is attached 'inside' the word, to its root or stem (depending on one's theoretical predilections), with concomitant loss or absence of the theme vowel. We also find two possible diminutive derivatives with athematic nouns like BrPt. *flor* 'flower': *florzinha*, where the attachment is to the whole word, and *florinha*, where it is attached 'word-internally'.[15] The minimal

[15] *Florzinha* is by far the more frequent of the two. *Florinha* occurs as a proper name in Brazil, though it is rare.

pair reveals that the epenthesis of /z/ does not take place at just any morpheme boundary, as Bachrach and Wagner originally observe, but only at the word level (however we encode that fact in our theory). Brazilian Portuguese /z/ epenthesis is thus shown to be morphologically complex in its conditioning, though in a different way from the other cases that we have discussed.

13.4 Other types of insertion

The factors conditioning the insertion of semantically empty segments given above can be extended to other types of insertion, including Catalan stem extenders (§13.4.1), the Italian augment <*isc*> (§13.4.2), and Spanish antesuffixes (§13.4.3). In each of these cases, the presence of the inserted syllable satisfies a phonological constraint, but its distribution is regulated by the morphology. We do not account for the segmental make-up of the inserted syllables.

13.4.1 Catalan stem extenders

In some dialects of Catalan, second-person singular imperatives of conjugation classes 2 and 3 have an 'extension' when enclitics are added. The 'extension' has no semantic role, but its presence is driven by a prosodic constraint: a moraic trochee is built at the right edge of the verb+enclitic unit (Bonet and Torres-Tamarit 2009).

(13) Formentera Catalan stem extenders

a.	[pɛ́ɾt] 'lose.2SG'	[pəɾð-ə̲́-lə] 'lose=it.F!'	([ə] extension)
b.	[əpɾə́n] 'learn.2SG'	[əpɾəŋ-gə̲́-lə] 'learn-it.F!'	([gə] extension)
c.	[búʎ] 'boil.2SG!'	[buʎ-i̯ɣ̲́-lə] 'boil=it.F!'	([iɣə] extension)

The form of the stem extender is taken from other forms of the imperative verb: it is the material that appears between the root and the person–number markers in the first- and second-plural imperatives. Compare the extensions in (13) with the first-person singular and plural forms in (14).

(14) Formentera Catalan first- and second-person plural imperatives

a.	[pɛ́ɾt] 'lose.2SG!'	[pəɾð-ə̲́-m] 'lose.1PL!'	(with [ə])
		[pəɾð-ə̲́-w] 'lose.2PL!'	
b.	[əpɾə́n] 'learn.2SG!'	[əpɾəŋ-gə̲́-m] 'learn.1PL!'	(with [gə])
		[əpɾəŋ-gə̲́-w] 'learn.+PL!'	
c.	[búʎ] 'boil.2SG!'	[buʎ-i̯ɣ̲́-m] 'boil.1PL!'	(with [iɣə])
		[buʎ-i̯ɣ̲́-w] 'boil.+PL!'	

While there is no consensus on the morphological status of this material, it is clear that its presence is driven by phonology, but its form is determined by morphology.

13.4.2 Italian <*isc*>

Italian has a meaningless augment, <*isc*>, with some fourth-conjugation-class verbs,[16] in the present indicative, present subjunctive, and imperative.[17] This is illustrated in (15) with the verb FINIRE 'finish'. (The forms are given in the orthography of Italian, except for stress, which is indicated here, though it is not in Italian orthography. <*isc*> before a back vowel is pronounced [*isk*], and before a front vowel is pronounced [iʃ].)

(15) *finire* 'to finish'

a. Present Indicative

fin-ísc-o	'1SG'	fin-iámo (**fin-isc-iámo)	'1PL'
fin-ísc-i	'2SG'	fin-íte (**fin-isc-íte)	'2PL'
fin-ísc-e	'3SG'		
fin-ísc-ono	'3PL'		

b. Present Subjunctive

fin-ísc-a	'1SG'	fin-iámo (**fin-isc-iámo)	'1PL'
fin-ísc-a	'2SG'	fin-iáte (**fin-isc-iáte)	'2PL'
fin-ísc-a	'3SG'		
fin-ísc-ano	'3PL'		

c. Imperative

		fin-iámo (**fin-isc-iámo)	'1PL'
fin-ísc-i	'2SG'	fin-íte (**fin-isc-íte)	'2PL'

Some argue that the presence of <*isc*> is phonologically motivated: it is only used with verb forms when it can be stressed, and is absent if it would be unstressed (Burzio and DiFabio 1994; Vogel 1994).[18] In other words, the augment regularizes the paradigm so that no verb forms are stressed on the stem.[19] In (16a) the <*isc*> augment is present between verb stem /fin/ and the first-person singular suffix /o/, and it is stressed; however, it is absent in (16b) because the stressed suffix /í/ would not allow stress to be realized on the augment.

(16) a. fin-ísc-o **fín-o 'I finish'
b. **fin-isc-í fin-í '(s)/he finished'

[16] Note that fourth-conjugation verbs with the augment are much more numerous than those without: 85% ~ 15% (Zamboni 1997:156).

[17] The etymology of the augment is to be found in Latin inchoative infix *-sc-* which was reanalysed in Italian and other Romance languages as a meaningless augment.

[18] Some claim that this fact about stress is a consequence of the presence of <*isc*> (Di Fabio 1990:4), and others, most notably Maiden (2004a:33) argue that 'the augment is not sensitive to stress; rather, stress and the augment are independently sensitive to the same, abstract, paradigmatic distribution'.

[19] In some Romance varieties the phonological conditioning is gone, and all forms have the augment: Nemi (Lazio) *feniscémo* 'finish.1PL', *feniscéte* 'finish.2PL' (Rohlfs 1968:243f.).

Hoberman and Aronoff (2003:74) show that the same augment found in Italian verbs is borrowed into Maltese (<*ixx*> in Maltese). It is found with the same verbs; however, the infix is distributed in different cells in the paradigms of the two languages. In Maltese, as in Italian, the augment is found 'in just those cases where the stem would otherwise be stressed'. The data in (17) show the distribution of the augment in the verb SUGGEST in the present and perfect indicative in Italian and Maltese.

(17)	Italian	Maltese
	sugger-ísc-o (**suggér-o)	ni-ssuġġer-íxx-i (**ni-suggér-i) 'I suggest'
		i-ssuġġer-í (**i-sugger-ixx-í) 'to suggest'
	sugger-í (**sugger-isc-í)	i-ssuġġer-íet (**i-ssuġġer-ixx-íet) 'she suggested'
		i-ssuġġer-íxx-a (**i-ssuġġér-a) 'he suggested'

Hoberman and Aronoff (2003:74) point out that 'this is a case of borrowing a phonological condition on a morphological rule'.

13.4.3 Spanish antesuffixes

Spanish antesuffixes or interfixes are inserted elements (<*ec*>, <*ol*>, etc.) whose distribution is morphologically conditioned (i.e., after certain stems and before certain derivational suffixes). These antesuffixes probably have the same etymology as the Brazilian Portuguese /z/. Latin had an *-ulus/-culus* alternation for diminutives (the latter formed from nouns ending in *-cus* + *-lus* diminutive). These elements have no meaning or connotative value (although in Italian they may; Prati 1942), but their distribution is phonologically influenced: short bases favour insertion of the antesuffix, as do bases ending in a sonorant consonant (except /l/) (Dressler and Merlini Barbaresi 1994).

(18)	Spanish		
	papel	'paper'	papel-ito '(DIM)'
	madr-e	'mother'	madr-ec-ita '(DIM)'
	comadr-e	'godmother'	comadr-ita '(DIM)'

The antesuffixes are productive, although unpredictable, so that the same base and suffix allow for various forms (Aguero-Bautista 1998), and the interfixed/non-interfixed form can have a special lexical meaning, or the two can be synonymous. Related antesuffixal interfixes can be found in other Romance languages as well: It. *boccon-c-ino* 'little bite' (Dressler and Merlini Barbaresi 1994; Napoli and Reynolds 1995). These are yet other examples of inserted segments that are sensitive to phonological structure, but their form and distribution are not the purview of phonology.

13.5 Findings

The factors conditioning the insertion of semantically empty material are given in Table 13.2. We see that in canonical epenthesis, the motivation for insertion is phonological, and morphology plays no role in the distribution and quality of the segment. In the cases we examined, phonology similarly is the motivation for insertion, but morphology plays a role in its quality and distribution.

Table 13.2 Types of epenthesis

Inserted Elements	**intrusive vowel**	**purely phonological epenthesis**	**morphologically conditioned epenthesis**	
			special epenthetic segment	**syllable insertion**
counts as segment(s)	–	+	+	+
presence is phonologically motivated	depends on one's definition	+	+	+
distribution is influenced by morphology (or morphosyntax)	–	–	+	+
quality is influenced by morphology	–	–	+	±

We have shown that insertion of semantically vacuous material (segment or syllable) lies along a cline from phonetic to phonological to morphological conditioning, as diagrammed in Figure 13.1.

Fig. 13.1 Factors conditioning the insertion of semantically vacuous material

The traditional view of epenthesis, going back to the first uses of the term, leads us to see it as a quintessentially phonological phenomenon. Research over the last several decades has uncovered its phonetic and even physiological roots in the classic notion of ease of articulation. In this contribution, we have used data from a number of Romance languages to show how epenthesis has found its way into what more traditionally might be called the grammar of these languages: individual

morphemes, more abstract purely morphological structures, and morphosyntax. What epenthesis retains throughout is its dependence on pure form. Even in cases where entire strings of segments are epenthesized, they have no meaning. If what characterizes human language above all else is its double articulation (Martinet 1949), then epenthesis provides striking demonstration of the persistence, power, and beauty of this property.