Book Section:
Wheeler, M.W. and O'Neill, P. (2022) Sandhi phenomena. In: Ledgeway, A. and Maiden, M., (eds.) Cambridge Handbook of Romance Linguistics. Cambridge Handbooks in Language and Linguistics. Cambridge University Press, pp. 209-233. ISBN: 9781108485791.

https://doi.org/10.1017/9781108580410.008

The Cambridge Handbook of Romance Linguistics

Sandhi phenomena

Max W. Wheeler and Paul O'Neill

0. Introduction

We take sandhi to be primarily a morphological phenomenon. Adaptation to different sandhi contexts gives rise to allomorphy (paradigmatic variation). Such adaptation generally reflects natural phonological processes which tend to reduce the markedness of sequences of phonological elements. We acknowledge that Romance also attests subphonemic alternations in sandhi environments, and we draw attention to some such cases, but they belong rather to phonology in general than to sandhi in particular. Probably the most striking feature of Romance sandhi alternations is the readiness with which they may become morphologized or lexicalized. This outcome may arise from subsequent sound change that makes the original motivated alternation opaque, or from levelling of allomorphic alternation that makes the distribution of allomorphs opaque. Occasionally, a morphologized/lexicalized alternation may be (partly) remotivated, as is famously the case with *rafforzamento fonosintattico* (§3.3.2) in standard Italian. But the phenomena of elision (§1.2) and *liaison* (§3.3.1) in modern French exemplify morphophonemic arbitrariness with very extensive incidence.

Words (including clitics) in Romance may begin with either a consonant (C) or a vowel (V), and may end with either a consonant or a vowel. So we find the following sandhi contexts: (i) V.#C, (ii) V.#V, (iii) C.#V, and (iv) C.#C. Of these, only (i) V.#C results in a sequence which is not to some degree phonologically marked. That is, V.#C is consistent with the least marked type of syllable structure CV.(CV.CV...) that all languages have —and V#C is not immune from phonological processes, though such will not be limited to the sandhi context (e.g. *gorgia toscana —la#casa* [la.hasa]; see. §00). We consider in turn each of the other basic types of word contact, and review what types of markedness-reduction processes each might be expected on general grounds to give rise to. One general observation to be made is that Romance does not favour insertion (epenthesis) to make sequences less marked. Nothing like r-insertion (*Cuba*[ɹ] *and America* in non-rhotic English) occurs in Romance to avoid hiatus. Cases are found of vowels inserted to avoid codas, but curiously only in Sardinia. In the Catalan of Alghero, epenthesis of [i] breaks up C.#C contacts not tolerated internally, so, e.g. *tot tapat* 'all covered' [ˌto.t i taˈpat], *cent voltes* 'a hundred times' [ˌsen.t i ˈvɔltas], *animals petits* 'small animals' [aniˌmal.ʦ i paˈtiʦ], but *diun coses* 'they say things' [ˌdiwŋ.ˈkɔzas], where [ŋ.k] is an acceptable internal syllable contact (Lloret & Jiménez 2007). In Sardinian, underlying word-final consonants are quite rare, and mostly inflectional; C.‖ is avoided by paragoge of a vowel that copies the vowel preceding the consonant. Several examples may be seen in the data in (37) below.

V.#V: Such a structure violates constraints penalizing hiatus. As well as the familiar Optimality Theory constraint ONSET ('syllables have onsets') we suggest a more specific constraint *HIATUS ('no V.V'), for reasons to be mentioned shortly. Broadly speaking, Romance languages deal with such structures in the same way: stressed vowels are not altered (FAITHFULNESS TO PROSODIC HEADS is active); an unstressed vowel is deleted next to (or absorbed into) another vowel of the same quality; shwa [ə] (or [œ] in French]) is deleted next to a vowel of any other quality; other unstressed vowel sequences become diphthongs (synizesis/synaeresis); hiatus is retained between two stressed vowels or between a stressed vowel and an unstressed vowel of different quality other than shwa. In many languages (French, Catalan, English), shwa is the vowel minimally opposed to zero, and liable to be elided provided adjacent consonants are syllabifiable. Avoidance of hiatus is sandhi is discussed in §1.

C.#V, as it stands, is a marked structure that violates two syllable structure constraints: NoCoda ('syllables do not have codas'), and Onset. Typically in Romance, this markedness is resolved by resyllabification —the word-final consonant becomes the onset of the following syllable. The consequent misalignment (.C#V) between word structure and syllable structure is tolerated. But in French, *liaison* without *enchaînement* is a possibility (Encrevé 1988); a word-final consonant is pronounced, but remains in coda position before an initial vowel. In several Romance languages, the consonant in a .C#V (or C.#V) structure may be liable to lenition. This is the case, for example, in those Romance languages that preserve Latin final <-*s*> and retain a voicing contrast in fricatives, namely, French, Occitan, Catalan, and Portuguese. In these languages lenition (voicing of fricatives) may apply to other word-final fricatives before a following vowel (or voiced continuant).

C.#C gives rise to more alternation in Romance than either of the other marked structures just mentioned. NoCoda is violated; there are good, grounded, typological reasons why a preconsonantal coda (C.C) is a 'worse', or more marked, structure than a coda otherwise (that is, C.‖). The phonetic features of a coda consonant are always less readily perceptible than those of an onset consonant —it is transition between consonant and vowel that gives the best perception cues. Before another consonant the features of a coda consonant tend to be masked by anticipatory co-articulation of that following consonant. Put another way, a preconsonantal coda consonant is phonologically weak. In Romance, a coda consonant may be subject to deletion, or to various kinds of weakening (and neutralization): assimilation, of place, manner, or laryngeal feature; shift along the sonority hierarchy stop–fricative–sonorant/approximant–glide. When the coda is itself complex, all of the coda elements are liable to weakening or deletion. A deleted coda may leave traces of its former character in the following word-initial onset. Hence *rafforzamento fonosintattico* in central and southern Italian varieties and Sardinian (§3.3.2). While some Romance languages retain a wide range of phonologically motivated allomorphy involving variation in word-final consonants before an initial consonant (Catalan: §3.1, South-Western Occitan §3.2), in others, the reflexes of this structure have become lexicalized —final consonant deletion or liaison in French (§3.3.1), lexicalized word-initial *rafforzamento fonosintattico* in Central and Southern Italo-Romance (§3.3.2).

‖#C: We briefly mention a couple of historical processes of initial epenthesis, (i) the insertion of a mid or high front vowel before an initial #sC cluster (Lat. *sperare* > Sp. *esperar*), and (ii) the insertion of /a/ before initial #[r]- in Gascon (Lat. *rem* > Gascon *arren* 'nothing') and Campidanese Sardinian (*radio* 'radio' [arˈraðiu]) (Sampson 2010). It has been claimed that such epenthesis originated in a strictly sandhi environment, namely, C#sC, C#r-, but Sampson argues that the historical evidence does not support such an idea (of generalization from a post-consonantal environment).

C#‖: Though, in the general case, as we have suggested, an utterance-final coda consonant has better perceptual cues than a preconsonantal coda, there are some sounds, or some alternations, where the utterance-final variant may, in fact, be weaker. Such is the case for [ɾ], where a tap without a transition to a following sound is quite difficult to detect, leading to loss in this environment sooner, or rather, than elsewhere (Catalan, Andalusian Spanish, Brazilian Portuguese, Occitan, to some degree in French (before the unconditioned change [ɾ] > [ʁ]); in both Catalan and French, loss of final /r/ is lexically conditioned, but is more widely found in polysyllables than in monosyllables. A parallel case is final /n/, or rather, its reflexes: Vn# > alternation Ṽ.n#V ~ ṼN.#C ~ Ṽ.#‖. In the utterance-internal environments, nasality, partly transferred to the preceding vowel, is supported by the retained presence of an

articulated nasal consonant. In utterance-final position, nasality, in the vowel alone, is liable to go unperceived, whence loss here sooner than elsewhere (Catalan, Lengadocian Occitan).[1]

1. Vowel sandhi

We illustrate vowel sandhi with two examples involving elision. In the first (§1.1), from Barcelona Catalan, the conditions governing elision, glide-formation (in the case of unstressed high vowels), or the retention of hiatus are all phonological or syntactic. In the second, from standard French (§1.2), elision, which at one time was phonologically motivated, has become opaque, through what are synchronically lexical exceptions.

1.1. Elison of [ə] in Barcelona Catalan.

In Barcelona Catalan there are three unstressed vowels [ə], [i] and [u]. In V.#V sandhi contexts, [ə] may be subject to elision, or diphthong formation with one of the other unstressed vowels (i.e. [əj], [əw], [jə], [wə]); [i] and [u] may be subject to deletion (adjacent to [i] and [u], respectively), or glide formation. Or, with each of them, hiatus may be retained. We focus here on the treatment of [ə], for brevity, and to make a more direct comparison with French. The data are largely from Recasens (1993), and the interpretation from Wheeler (2005, chap. 4).

1.1.1 Stressed vowel followed by an unstressed non-high vowel [ə]

In eastern continental Catalan that is primarily considered here, as a result of vowel reduction in unstressed syllables, the only unstressed non-high vowel in initial position is [ə], spelt <*a*> or <*e*>. The possibilities in the case of input ˈV##ə are hiatus or elision. When the lowest, most prominent, vowel [ˈa] is followed by a word beginning with [ə], elision is obligatory (1a). With the other low vowels [ˈɛ] and [ˈɔ], marginally less prominent than [ˈa], elision also seems to be practically obligatory (1b).

(1) a. *està atordit* [əs.ˈta .tuɾ.ˈðit] ‘is stunned’
mesurar alçària [mə.zu.ˈɾa l.ˈsa.ɾjə] ‘measure.INF height’

b. *cafè amarg* [kə.ˈfɛ. ˈmark] ‘bitter coffee’
això anima [ə.ˈʃɔ. ˈni.mə] ‘that encourages’

What appears to be problematic about the realization of [ˈa.ə], [ˈɛ.ə] and [ˈɔ.ə] is that a stressed vowel of high perceptual prominence is followed by the unstressed vowel of minimum prominence [ə]. Perceptually it is as if the prominent vowel overwhelms the non-prominent vowel, or obscures its presence entirely. Because such sequences are hard to perceive, speakers may not take the trouble to make the sequence at all, despite loss of semantic information in some cases.

In the case of the less prominent stressed vowels followed by [ə], however, other considerations may come into play, favouring hiatus.

(2) *actor enèrgic* [ək.ˈto .ə.ˈnɛr.ʒik] ‘energetic actor’
no animes [ˈno .ə.ˈni.məs] ‘not encourage.2SG.PRS.IND’.

(3) *actor esplèndid* [ək.ˈtos.ˈplɛn.dit] ‘splendid actor’
collir espàrrecs [ku.ˈʎis.ˈpa.rəks] ‘pick.INF asparagus’

In (2), elision of [ə] in an open syllable would lead to a stress clash e.g. *actor enèrgic* [ək.ˈto .ˈnɛr.ʒik] —more particularly, a clash of adjacent stressed moras. In (3), however,

[1] We use *utterance* here with deliberate vagueness. Though the processes in question always apply at the start or end of an utterance, as the case may be, they may also be found at the edge of an intonational phrase, or a phonological phrase, in ways that vary between and within languages (or between varieties, or speakers).

with [ə] in a closed syllable, clash of adjacent stressed syllables with a mora between the heads, is preferable to hiatus. Other considerations, too, may favour or disfavour hiatus.

(4) a. *corder anyal* [kuɾ.ˈðe .ə.ˈɲal] ‘lamb less than one year old’, cf. *any* ‘year’
setí arnat [sə.ˈti. ər.ˈnat] ‘moth-eaten satin’, cf. *arna* ‘moth’

b. *sentir avidesa* [sən.ˈti .ə.βi.ˈðɛ.zə] ‘feel.INF greed’, cf. *àvid* ‘greedy’

c. *actor alpinista* [ək.ˈtol.pi.ˈnis.tə] ‘mountaineering actor’, cf. *Alp* ‘Alp’

d. *tabú acceptable* [tə.ˈβuk.səp.ˈtab.blə] ‘acceptable taboo’
tauló allargat [təw.ˈlo.ʎər.ˈɣat] rather than *[təw.ˈlo .ə.ʎər.ˈɣat] ‘long plank’

In (4a-b), a word-initial [ə] corresponds to a stressed vowel in the base word. Elision is disfavoured. However, despite the stressed base *Alp* in (4c), elision takes place because *[ək.ˈto. əl.pi.ˈnis.tə] would fall foul of a LAPSE constraint against more than two moras intervening between the head syllables of prosodic words. (CLASH and LAPSE constraints together favour stress on alternate syllables.) Finally, hiatus is more likely after a monosyllabic major lexical item (5). Though elision in (5b) shows that elision is reckoned better than a three-syllable lapse between stresses.

(5) a. *fi estranya* [ˈfi .əs.ˈtɾa.ɲə] ‘unusual end’
té animals [ˈte .ə.ni.ˈmals] ‘has animals’
vi exquisit [ˈbi .əks.ki.ˈzit] ‘select wine’

b. *vi extraordinari* [ˈbiks.tɾəwr.ði.ˈna.ɾi] ‘extraordinary wine’

1.1.2. Unstressed non-high vowel [ə] *followed by a stressed vowel*

When a non-high unstressed vowel ([ə]) precedes a stressed one, as before, elision is blocked when a clash (one mora or none intervening) between phonological phrase heads would result, as in the examples of (6).

(6) *foca àrtica* [ˈfo.kə .ˈar.ti.kə] ‘arctic seal’ (*foca*]N *àrtica*]A)

rega arbres del jardí [ˈrɛ.ɣə .ˈa.βɾəz .ðəl .ʒər.ˈði] ‘water.3SG.PRS.IND trees in the garden’ (*rega*]V *arbres*]N *del jardí*]PP)

assenyalar-ne una [ə.sə.ɲə.ˈlar.nə .ˈu.nə] ‘point-out.INF one of-them’

Note that whereas hiatus between a preceding stressed low vowel and [ə] is highly disfavoured (1), when [ə] is *followed* by a stressed low vowel, hiatus is preferred to a stress clash.

When the first word is in specifier position, or is a preposition (thus not a phonological phrase head and not protected by CLASH constraints), elision is normal (7), reflecting the domination of ONSET over faithfulness constraints.

(7) *la primera hora* [lə .pɾi.ˈme.ˈɾɔ.ɾə] ‘the first hour’
onze anys [ˈon.ˈzaɲs] ‘eleven years’
entre altres [ˈen.ˈtɾal.tɾəs] ‘among others’

1.1.3. Contact between unstressed vowels

Having dealt with sequences of vowels one of which is stressed, we turn to contact between unstressed vowels. When unstressed vowels come into contact, the CLASH constraint is not relevant. Reducing two unstressed syllables to one unstressed syllable can never lead to a CLASH violation. However, reducing two unstressed syllables to one would often lead to a more harmonic outcome. In fact, when each of the adjacent vowels is [ə], elision/fusion is the regular result, as in *perdre alè* [ˈpɛr.ðɾə.ˈlɛ] ‘lose.INF breath’, *escriptora aguda* [əs.kɾip.ˈto.ɾə.ˈɣu.ðə] ‘penetrating writer.F’, with elision/fusion reflecting the favouring of the anti-hiatus constraints over faithfulness.

When an unstressed high vowel ([i] or [u]) precedes an unstressed non-high vowel ([ə]), three possible outcomes need to be considered: hiatus ([i.ə], [u.ə]), a rising diphthong ([jə], [wə]), and elision of the non-high vowel [ə]. For each case, often two of the three alternatives occur in variation, while the third is ungrammatical. Where hiatus and a rising diphthong are both acceptable, hiatus generally belongs to a more formal or conservative style of pronunciation (reflecting relatively high ranking of constraints disfavouring glides in complex onsets as also word-internally). When a non-high unstressed vowel ([ə]) precedes a high one, a falling diphthong ([əj], [əw]) or elision of [ə] are both possible. Bonet & Lloret (1998: 185) remark that elision is commoner, for example, in the Girona region than in the Barcelona region.

1.2. Elision in French

In Old and Middle French, word-final (unstressed) [ə] (later [œ]) was elided before a vowel-initial word, avoiding hiatus; the final [a] of the feminine definite article *la*, and the homophonous feminine singular direct object pronoun, were treated in the same way. In the case of the articles/pronouns, *le*, *la*; the personal pronouns *je*, *se*, *me*, *te*; the preposition *de*; the conjunction *que*, the elision was, and is, represented with an apostrophe: *le* + *arbre* → *l'arbre* 'the tree', *la* + *étoile* → *l'étoile* 'the star', *me* + *embêter* → *m'embêter* 'annoy me'; but elision will also always have occurred in e.g. *quelqu*ɐ́ *obstacle* 'some obstacle'. No issue would have arisen, had not word-initial /h/ been lost in regular sound change, without the subsequent vowel-initial reflexes triggering the expected elision of preceding [ə] (Burov 2012: 187-192; Tranel 1987: 93-96; 228-229).

(8) *le hibou* [lœ .i.bu] 'the owl'
la hauteur [la .o.tœr] 'the height'
je hais [ʒœ .a.i] 'I hate'
se heurter [sœ .œr.te] 'to crash'
quelque honte [kɛl.kœ .õt] 'whatever shame'

At this stage of the language, elision has become lexicalized: it applies before most vowel initial words, but not all. Hiatus after [œ] is no longer ungrammatical. Words that etymologically had no /h-/ may be added to the exception list, e.g. *le onze août* [lœ .õn.z ut] 'the 11th of August' (*onze* < UNDECIM); *le héros* [lœ .ero] (where the <*h*-> was merely orthographic already in Latin). As well as under-application of elision making it an opaque process, over-application does likewise: elision takes place in the case of most, but not all, words that begin with a glide (9).

(9) a. /j/ *n'avoir d'yeux que pour* [na.vwar. djø .kœ .pur] 'to only have eyes for'
le match d'hier [lœ. matʃ. djɛr] 'yesterday's match'
l'iode [ljɔd] 'iodine'

b. /w/ *l'oiseau* [lwa.zo] 'the bird'
l'ouate [lwat] 'the cotton wool'
l'ouïe [lwi] 'hearing'
le vent d'ouest [lœ vɛ̃. dwɛst] 'the west wind'
l'oindre [lwɛ̃dr] 'to anoint him'

c. /ɥ/ *l'huile* [lɥil] 'oil'
l'huître [lɥitr] 'the oyster'

Elision before some of the word types in (9) has a historical justification, in that they originally began with a falling, not a rising diphthong: /oe/ in *oiseau*, /oĩ/ in *oindre,* /yj/ in *huile*; or with syllabic vowels in hiatus: /u.i/ in *ouïe*. Others, though, are likely to have had a rising diphthong as soon as they had a diphthong at all: *yeux, hier*. Or, a different interpretation might be that, at one stage of French, the sonority gradient from a glide to a

vowel was not sufficient to count as a syllable onset. In any case, however, other glide-initial words, some but not all of which are recent borrowings, do not trigger elision (10):

(10) a. /j/ *le yaourt* [lœ .ja.urt] 'yoghurt'
le yogi [lœ .jɔ.gi] 'the yogi

b. /w/ *le watt* [lœ .wat] 'the watt' (NB. minimal pair with *l'ouate* in (9)).
le whisky [lœ .wis.ki] 'whisky'

c. /ɥ/ *le huit* [lœ .ɥit] 'the 8th'
pour le huer [pur .lœ .ɥe] 'in order to boo him'

Both the vowel-initial words (8) and the glide-initial words (10) that reject elision are at least consistent in being treated as if they began with an onset consonant in other syntactic contexts in addition to these, namely, in rejecting liaison (§3.3.1), and in selecting the preconsonantal allomorphs of those words that have suppletive forms: *ce* rather than *cet* 'this', *les* [le] rather than [lez] 'the.PL', *du* rather than *de l'* 'of the', *vieux* rather than *vieil* 'old', etc. Nevertheless, 'being treated as if they began with a consonant' does not extend to rejecting resyllabification of a preceding word-final consonant as an onset ('*enchaînement*') (11).

(11) *le hasard* 'chance' but *par hasard* [pa .r a.zar] 'by chance'
le hibou 'the owl', but *quel hibou* [kɛ .l i.bu] 'what owl'
le onze 'the 11th', but *deux mille onze* [dø .mi .l ɔ̃z] '2011'

The morphologization/lexicalization of elision is French is the more remarkable as the original distribution was so simply motivated on a straightforward markedness principle, and as its effect is so pervasive in incidence, in text and in the lexicon. NB. in those varieties of French in which by general rule all final [œ] is lost except when the surrounding consonants cannot otherwise be syllabified, [œ] is retained before the anomalous initial vowels (8) and glides (10).

2. Inter-word vowel-consonant contact — V.#C:

The sandhi context V.#C is as unmarked as possible from the point of view of syllable structure. However, some Romance languages have an active lenition process in which initial consonants are treated the same as medial ones. Thus, in Galician, Spanish (but not Judeo-Spanish (Hualde, 2013:243-244), Catalan and south-western Occitan (Gascon and Lengadocian), voiced non-strident obstruents /b/, /d/, /g/ are pronounced as fricatives or approximants ([β], [ð], [ɣ]) between continuants (that is, vowels, glides, /r/, /l/, /ʎ/, /z/, /ʒ/), and pronounced as stops elsewhere. This areal phenomenon is shared with Basque.

(12) Sp. *las bodas* [laz βoðas] 'the weddings'
la dicha [la ðiʧa] 'happiness'
el gordo [el ɣoɾðo] 'the fat one'

The distribution of lenition here is markedly different from the earlier process affecting Latin obstruents in the transition to Western Romance, when initial consonants were always preserved. The same kind of active lenition of /b/, /d/, /g/ is found also in Sardinian (see examples in (37)), and there are comparable phenomena in several south Italian dialects; e.g. in Neapolitan /b/ [b] ~ [v], /d/ [d] ~ [ɾ], /g/ [g] ~ [ɣ] (Ledgeway 2009: 39). In Tuscan, also, while voiced obstruents may be variably lenited (with results as in South-Western Romance just mentioned), voiceless stops are also lenited ('spirantized') in such contexts (13) (*gorgia toscana*, cf. Sorianello 2010).

(13) Tuscan *la patata* [la ɸaˈθaθa] ‘the potato’
la piega [la ˈɸjɛɣa] ‘the fold’
la tavola [la ˈθavola] ‘the table’
la crema [la ˈhrɛma] ‘cream’

In South-Western Romance the lenition alternations of voiced obstruents are automatic and sub-phonemic, and have no morphological consequences. The same would be true in the Italian and Sardinian varieties were it not for the fact that these varieties also attest *rafforzamento fonosintattico* (RF —see below §3.3.2). In RF initial consonants are strengthened —typically geminated— after a lexical set of vowel-final words, a set which does not include the feminine singular definite article *la*. So we can find partly lexically conditioned morphological alternations such as (14).

(14) Tuscan: *in casa* [iŋ ˈkasa] ‘at home’ – *la casa* [la ˈhasa] ‘the house’ – *da casa* [da kˈkasa] ‘from home’.

We consider this phenomenon (RF) more fully in the context of consonantal sandhi contacts C.#C.

3. Inter-word consonant-consonant contact — C.#C

Only certain Romance languages have a large proportion of consonant final words with a large variety of consonants. In standard Italian, and many Italian dialects (Ligurian, Venetian, and all those of the centre and South), few words are consonant-final. In Spanish and Portuguese there are somewhat more, though many consonants (obstruent stops, and labials) are excluded from final position. In the remainder, a wide range of word-final consonants occur.

3.1. Consonant contacts in Majorcan Catalan

We consider first a case where allomorphic alternation in consonant-final words before initial consonants is both regular and extensive. The case is that of Majorcan Catalan, where a consonant of almost any place and manner can end a word; only voicing contrast is neutralized in word-final position before a vowel or a pause (= citation form) (Bibiloni 2016: 152-169; Wheeler 2005: 207-249 with data from Dols 1993). First, final plosives and nasals are subject to place assimilation before a following consonant (15). (For /ɲ/ see below.)

(15) Majorcan place (and voice) assimilation

a. /p/ + /t/: *cap tros* ‘no chunk’ [ˌkat. ˈtɾɔs]
/p/ + /d/: *cap dit* ‘no finger’ [ˌkab. ˈdit]
/p/ + /g/: *cap goma* ‘no rubber’ [ˌkag. ˈgomə]
/t/ + /p/: *set parts* ‘seven parts’ [ˌsɛp. ˈpaɾs]
/t/ + /b/: *set braços* ‘seven arms’ [ˌsɛb. ˈbɾasos]
/t/ + /k/: *set cases* ‘seven houses’ [ˌsɛkˈ kazəs]
/t/ + /g/: *set gàbies* ‘seven cages’ [ˌsɛg. ˈgaβis]
/k/ + /p/: *puc passar* ‘I can go by’ [ˌpup. pəˈsa]
/k/ + /b/: *puc beure* ‘I can drink’ [ˌpub. ˈbəwɾə]
/k/ + /t/: *puc treure* ‘I can take out’ [ˌput. ˈtɾɛwɾə]
/k/ + /d/: *puc dormir* ‘I can sleep’ [ˌpud. doɾˈmi]

b. /p/ + /s/: *cap sac* ‘no bag’ [ˌka. ˈʦak]
/p/ + /z/ *cap zero* ‘no zero’ [ˌka ˈʣero]
/p/ + /ʃ/ *cap xeringa* ‘no syringe’ [ˌka. ʧəˈɾiŋgə]
/p/ + /ʒ/ *cap jardí* ‘no garden’ [ˌka. ʤəɾˈði]
/k/ + /s/ *puc sortir* ‘I can go out’ [ˌpu. ʦoɾˈti]
/k/ + /z/ *duc zinc* ‘I bring zinc’ [ˌdu. ˈʣiŋk]

/k/ + /ʃ/ *puc xerrar* ‘I can talk’ [ˌpu. ʧəˈra]
/k/ + /ʒ/ *puc jugar* ‘I can play’ [ˌpu. ʤuˈɣa]

c. /n/ + /p/ *són petits* ‘they are small’ [ˌsom. pəˈtiʦ] (= *som petits* ‘we are small’)
/n/ + /b/ *són bons* ‘they are good’ [ˌsom. ˈbɔns] (= *som bons*)
/n/ + /m/ *són molts* ‘they are many’ [ˌsom. ˈmols] (= *som molts*)
/m/ + /t/ *som tots* ‘we all are’ [ˌson. ˈtoʦ] (= *són tots*)
/m/ + /d/ *som dos* ‘we are two’ [ˌson. ˈdos] (= *són dos*)
/m/ + /k/ *som quatre* ‘we are four’ [ˌsoŋ. ˈkwatɾə] (= *són quatre*)
/m/ + /g/ *som grans* ‘we are grown up’ [ˌsoŋ. ˈgɾans] (= *són grans*)
/m/ + /f/ *som feliços* ‘we are happy’ [ˌsoɱ. fəˈlisos] (= *són feliços*)
/m/ + /v/ *som vius* ‘we are alive’ [ˌsoɱ. ˈviws] (= *són vius*)
/m/ + /s/ *som set* ‘we are seven’ [ˌson. sɛt]
/m/ + /l/ *som liberals* ‘we are liberal’ [ˌson. liβəˈɾals]
/m/ + /ʎ/ *som llests* ‘we are clever’ [ˌsonʲ. ʎeʦ]
/m/ + /n/ *som nous* ‘we are new’ [ˌson. ˈnɔws]

Obstruent stops and labiodental fricatives before labiodentals and sonorants are subject to place and manner assimilation (16). In (16c) the alveolar trill [r] is already reckoned a geminate, so is not duplicated in the transcription. In (16e) a final /f/, before a consonant, behaves just like a final /t/; only a selection of examples is given

(16) a. /p/ + /f/ *cap flor* ‘no flower’ [ˌkaf. ˈflɔ]
/p/ + /v/ *cap vidre* ‘no glass’ [ˌkav. viðɾə]
/t/ + /f// *set forats* ‘seven holes’ [ˌsɛf. foɾaʦ]
/t/ + /v/ *set vots* ‘seven votes’ [ˌsɛv. vɔʦ]
/k/ + /f/ *sac foradat* ‘bag with holes’ [ˌsaf. foɾəˈðat]
/k/ + /v/ *puc venir* ‘I can come’ [ˌpuv. vəˈni]

b. /p/ + /l/ *cap licor* ‘no liqueur’ [ˌkal. liˈkoɾ]
/p/ + /ʎ/ *cap lladre* ‘no thief’ [ˌkaʎ. ˈʎaðɾə]
/t/ + /l/ *set làmines* ‘seven plates’ [ˌsɛl. ˈlaminəs]
/t/ + /ʎ/ *set lladres* ‘seven thieves’ [ˌsɛʎ. ˈʎaðɾəs]
/k/ + /l/ *puc lamentar* ‘I can complain’ [ˌpul. ləmənˈta]
/k/ + /ʎ/ *puc llegir* ‘I can read’ [ˌpuʎ. ʎəˈʤi]

c. /p/ + /r/ *cap rata* ‘no rat’ [ˌka. ˈratə]
/t/ + /r/ *set rates* ‘seven rats’ [ˌsɛ. ˈratəs]
/k/ + /r/ *puc riure* ‘I can laugh’ [ˌpu. ˈriwɾə]

d. /p/ + /m/ *cap mà* ‘no hand’ [ˌkam. ˈma]
/p/ + /n/ *cap nin* ‘no boy’ [ˌkan. ˈnin]
/t/ + /m/ *set mans* ‘seven hands’ [ˌsɛm. ˈmans]
/t/ + /n/ *set nins* ‘seven boys’ [ˌsɛn. ˈnins]
/k/ + /m/ *puc mirar* ‘I can look’ [ˌpum. miˈɾa]
/k/ + /n/ *poc net* ‘not very clean’ [ˌpɔn. ˈnət]

e. /f/ + /p/ *agaf pomes* ‘I pick apples’ [əˌɣap. ˈpoməs]
/f/ + /t/ *agaf taronges* ‘I pick oranges’ [əˌɣat. təˈɾɔɲʒəs]
/f/ + /ʒ/ *agaf gelat* ‘I choose ice cream’ [əˌɣa. ʤəˈlat]
/f/ + /r/ *agaf roses* ‘I pick roses’ [əˌɣa. ˈrɔzəs]
/f/ + /m/ *agaf mores* ‘I pick blackberries’ [əˌɣam. ˈmoɾəs]
/f/ + /n/ *agaf nesples* ‘I pick medlars’ [əˌɣan. ˈnespləs]

There are some slightly unusual modifications in the case of word-final nasals. Before an alveolo-palatal sibilant, as well as assimilating place, /m/ and /n/ trigger affrication (17), a

process also found word-internally. As seen in (15c), after a nasal there is no affrication of an alveolar sibilant.

(17) /m/ + /ʃ/ *hem xerrat* 'we have talked' [ˌənʲ. ʧəˈrat]
/m/ + /ʒ/ *som joves* 'we are young' [ˌsonʲ. ˈʤovəs] (= *són joves*)
/n/ + /ʒ/ *en Joan* 'Joan'[2] [ənʲ. ʤuˈan]

In preconsonantal position the palatal nasal /ɲ/ splits into a palatal glide and a nasal that assimilates place (18).

(18) /ɲ/ + /d/ *lluny de tu* 'far from you' [ˌʎujn. də ˈtu]
/ɲ/ + /p/ *l'any passat* 'last year' [ˌlajm. pəˈsat]
/ɲ/ + /k/ *any curt* 'short year' [ˌajŋ. ˈkuɾt]

Internally, the same process is found before inflectional /+s/, e.g. *estrenys* 'you.SG squeeze' [əsˈtrəjns], which is anomalous in displaying a coda with three consonants, when Majorcan otherwise has a strict maximum of two consonants in a coda, so *camps* 'fields' (/kamp/ + /s/) is [ˈkans].

A final alveolo-palatal obstruent becomes a palatal glide before an initial consonant (19):

(19) /ʃ/ + C *peix frit* 'fried fish' [ˌpej. ˈfrit], *calaix petit* 'small drawer' [kəˌlaj. pəˈtit]
/ʧ/ + C *vaig dir* 'I said' [ˌvaj. ˈði], *Puig Major* 'Great Hill' [ˌpuj. məˈʒo].

Final /s/ undergoes total assimilation before a liquid or a (palatal) glide (20a). Alveolar /s/ plus a sibilant merge as an affricate with the place of the onset consonant (20b). Before other voiced consonants /s/ becomes [ɾ] (20c). Before other voiceless consonants, /s/ is preserved.

(20) a. /s/ + /l/ *és logic* 'it is logical' [ˌəl. ˈlɔʒik]
/s/ + /ʎ/ *es llit* 'the bed' [ˌəʎ. ˈʎit]
/s/ + /r/ *ses rates* 'the rats' [ˌsə. ˈratəs]
/s/ + /j/ *es iogurt* 'the yoghurt' [ˌəj. joˈɣuɾt

b. /s/ + /s/ *les set* 'seven o'clock' [lə. ˈʦɛt]
/s/ + /ʃ/ *es xat* 'the chat' [ə. ˈʧat]
/s/ + /ʒ/ *dos joves* 'two young people' [ˌdo ˈʤovəs]

c. /s/ + /b/ *és bo* 'it is good' [ˌəɾ. bɔ]
/s/ + /d/ *és dolent* 'it is bad' [ˌəɾ. ðoˈlent]
/s/ + /v/ *dus vi* 'you bring wine' [ˌduɾ. ˈvi]
/s/ + /m/ *és meu* 'it is mine' [ˌəɾ. mew]

Final liquids are largely resistant to modification before initial consonants, though /ɾ/ variably assimilates to a following lateral: *per la casa* 'through the house': [pəl. lə ˈkazə] ~ [pəɾ. lə ˈkazə], *per llogar* 'for hire' [pəʎ. ʎoˈɣa] ~ [pəɾ. ʎoˈɣa].

Just to spell out again the degree of allomorphy these processes give rise to, *set* 'seven', for example, has the following 12 variants in the data given above: [sɛt, sɛp, sɛb, sɛk, sɛg, sɛf, sɛv, sɛl, sɛʎ, sɛr, sɛm, sɛn].

The Majorcan examples we have considered so far in this section have just one consonant in the word-final coda. Majorcan Catalan words can readily have complex codas with two consonants (not more), which are pronounced in phrase-final position, and before vowel-initial words, where the second coda consonant is resyllabified as an onset. Before a consonant-initial word, however, a coda must consist of at most one consonant.

That is, in Majorcan, a pre-onset coda cluster constraint *CC]σC is active. The motivation is fundamentally perceptual. There are inadequate cues to the place and/or manner of a

[2] *en* is the 'personal article' used before MASC.SG. proper names.

consonant, especially a stop, between two consonants. The context is somewhat less favourable than in a phrase-final CC cluster. Phrase-final clusters of CC]σ form are likely to be pronounced more emphatically, since phrase-final position is also that of nuclear stress, so there is likely to be more time for the speaker to achieve articulatory targets, and phrase-final position also allows more time for perception of the relatively weakly cued elements than in phrase-internal position.

The effect of this constraint in Majorcan is that deletion of a consonant from a coda cluster is quite extensive. In the general case, it is the middle consonant of the cluster CC.#C, the word-final one that is deleted. Quite often this consonant realizes an inflectional morpheme. The remaining consonant is subject to the assimilation processes mentioned above. The following examples (21) are taken in the main, again, from Bibiloni (2016).

(21) *plats grocs* 'yellow plates' [ˌplaɡ. ˈgɾɔʦ]
llocs deshabitats 'uninhabited places' [ˌʎɔd. dəzəβiˈtaʦ]
obr sa porta 'I open the door' [ˌɔ. ʦə ˈpɔɾtə]
ho arregl tot 'I arrange it all' [ə.w əˌret. ˈtot]
resolc problemes 'I solve problems' [rəˌzɔl. pɾobˈbleməs]
molts d'anys 'many years' [ˌmol. ˈdajns]. NB here two underlying consonants are deleted, but *molts* is anyway [mols] in phrase-final position, by the general *CCC]σ constraint.
ells canten 'they are singing' [ˌeʎ. ˈkantən]
porc negre 'black pig' [ˈpɔɾ. ˈnəɣɾə], cf. *porcs negres* 'black pigs' [ˈpɔɾ. ˈnəɣɾəs]
temps de figues 'fig season' [ˌten. də ˈfiɣəs]
ponts baixos 'low bridges' [ˌpɔm. ˈbaʃos]
pens que sí 'I think so' [ˌpəŋ. kə ˈsi]
lluit sense por 'I fight fearlessly' [ˌʎuj. ˌsənsə ˈpɔ]

As in (18), /ɲ/ undergoes an unusual split so that both height and nasality are retained but sequenced (22a). The clusters /nk/[3] and /nʃ/ (/nʒ/) are treated in the same way (22b).

(22) a. *anys difícils* 'difficult years' [ˌajn. diˈfisils]
punys bruts 'dirty fists' [ˌpujm. ˈbɾuʦ]

b. *banc blau* 'blue seat' [ˌbajm ˈblaw]
menj mel 'I eat honey' [ˌməjm. ˈmɛl]

Whereas generally the outside consonant in a word-final cluster is lost, atypically, in the case of /-sC-/ clusters, and /-ɾn-/, the prefinal consonant goes. So we get as in (23).

(23) *he vist barques* 'I have seen boats' [e ˌvib ˈbaɾkəs]
bosc cremat 'burnt wood' [bɔk. kɾəˈmat]; *boscs cremats* 'burnt woods' [bɔk. kɾəˈmaʦ]
aquest nin 'this boy' [əˌken. ˈnin]
gusts nous 'new flavours' [ˌgun. ˈnɔws]
carn freda 'cold meat' [ˌkaŋ. fɾəðə]; *carns fredes* 'cold meats' [ˌkaŋ. fɾəðəs]
forn calent 'hot oven' [ˌfoŋ. kəˈlent]

[3] Nicolau Dols and Gabriel Bibiloni (personal communication) inform me that /nk/ is not treated in this way where /k/ realizes an allomorph of 1SG.PRS.IND; here /k/ is simply deleted preconsonantally: *prenc nota* 'I take note' [ˌpɾən.ˈnɔtə], *venc llibres* 'I sell books' [ˌvənʲ. ˈʎiβɾəs]. As far as I am aware, this is the only case in which a morphological consideration affects word-final consonant sandhi in Majorcan.

3.2. Consonantal contact in Occitan

The treatment of C.#C sandhi in South-western Occitan (Gascon and western Lengadocian) is broadly similar to that in Majorcan Catalan (Burov 2012: 204, 211, 230; Loporcaro 1997: 73fn.). Here are some examples (24a-b).

(24) a. Aranese Gascon

poc pan 'not much bread' [ˌpɔp. paŋ]
eth pè 'the foot' [ɛp. ˈpɛ]
eth mur 'the wall' [ɛm. ˈmyɾ]

b. western Lengadocian

lop gris 'grey wolf' [ˌlug. ˈgɾis]
sèrp d'aiga 'water snake' [ˌsɛɾ.ˈðajɣɔ]; cf. *sèrps d'aiga* 'water snakes' [ˌsɛɾ.ˈðajɣɔ]
avèm pas de clients per la vendre 'we haven't clients to sell it to' [aˌβɛm. pa ðe kliˈem. peɾ la ˈβendɾe]
mieg nut 'half naked' [ˌmjɛn. ˈnyt]
sap faire 'knows how to do' [ˌsaf. ˈfajɾe]
taps longs 'long corks' [ˌtal. ˈluns]
aquèl cat blanc dormís 'that white cat is asleep' [aˌkɛl. ˌkab. ˈblan. duɾˈmis]
un còp me venguèt quèrre 'once she came to look for me' [yŋ.ˈkɔm. me βeŋˈgɛk. ˈkɛre]

A characteristic of eastern Lengadocian, shared with Provençal, is the simplification of geminates, internally, and across boundaries. Hence an eastern Lengadocian variant of the last example of (24) would be [yŋ.ˈkɔ. me βeŋˈgɛ. ˈkɛre]. A significant consequence of this degemination process is that it derives allomorphs in which a word-final consonant is absent altogether. This is plausibly one of the routes by which other Occitan varieties, principally, Provençal, Lemosin, and Auvernhat, have lost all word-final obstruents, though Provençal retains /s/ after a stressed vowel and also inflectional /s/ marking 2nd person singular, but not /s/ marking nominal plurals.

3.3. Lexicalization in C.#C contacts

In the previous sections (§3.1, §3.2) we have described situations in which deletion in word-final codas is triggered by phonological context — position before a consonant-initial word. We have not focused on the variants found in the other contexts, taking it for granted that, as is very generally the case, no deletion occurs there, though we could point out that, in Catalan and Occitan, voicing contrasts in final obstruents are neutralized everywhere: before vowels, as a rule, stops are voiceless and fricatives are voiced;[4] in phrase-final position, all obstruents are voiceless. We contrast this state of affairs with two cases in which original alternations in final-consonant words have become opaque, and lexicalized.

3.3.1 French *liaison*

In French, *liaison* is the name given to the phenomenon whereby the final consonants (as represented by the orthography) of certain words are pronounced before vowel-initial words in certain contexts but not elsewhere. Which consonants, in which words, and in which contexts, is determined only to a small degree by phonological factors; the determining features are lexical, and partly morphological or syntactic. In this section, we attempt to give

[4] The voicing of final (strident) fricatives before vowels is common to the western Romance languages that display word-final fricatives and have a voicing contrast in fricatives, i.e. Portuguese, Catalan, Occitan, French, Rheto-Romance, North Italian dialects, and Sardinian. Spanish and Galician will have shared this before merging voiced fricatives with voiceless generally.

some account of how this situation came about. The original distribution —of alternants with and without final consonants— was determined, it is believed, by phonological context alone.

In Old French, a word could end in any one consonant. A word could end in two consonants if the first was a nasal or /r/.[5] Modern French orthography generally reflects this state of the language: *champ* 'field', *plomb* 'lead', *font* 'source', *long* 'long', *vend* 'sells', though in modern French these words are open syllables [ʃɑ̃], [plɔ̃], [fɔ̃], [lɔ̃], [vɑ̃] (unusually a final consonant is retained in *cinq* 'five' [sɛ̃k], and *donc* 'thus' [dɔ̃k]); *faubourg* 'suburb' [fobuʀ], *vert* 'green' [vɛʀ], *part* 'part' [paʀ], *porc* 'pig' [pɔʀ], *cerf* 'stag' [sɛʀ], *sers* 'serve.2SG.PRS.IND', (unusually both consonants are retained in *mars* 'March' [maʀs], *ours* 'bear' [uʀs], *parc* 'park' [paʀk], and variably *serf* 'serf' [sɛʀ(f)]). Morin (1986: 168) proposes that, by late Old French, word-final clusters were permitted only in phrase-final position, not before vowels. Between the 12th and the 16th centuries, final consonants progressively disappeared, first in the weakest, preconsonantal coda position, then, generally, in phrase-final codas. Later, the range of contexts in which the remaining prevocalic alternants appeared began to shrink, giving the pattern of modern French liaison, grammatically and lexically conditioned. Morin (1986: 169) suggests that preconsonantal /s/ weakened to /h/, then remained as vowel lengthening, before disappearing altogether. This would parallel the historical treatment of preconsonantal /s/ within words.

A table of the allomorphs of consonant-final numerals in modern French (25) gives some idea of what the distribution may have been more generally in consonant-final words once a rule was introduced deleting a coda in preconsonantal position. Even here, though, there has been a considerable amount of redistribution of allomorphs, short of complete levelling. We take it that the earlier regular distribution is that which can be observed for *un* '1', *six* '6', *huit* '8' and *dix* '10', that is, the final consonant is pronounced before a vowel (liaison context), and, except for /n/ (*un*), also in phrase-final position. The current distribution of allomorphs displays complex and inconsistent trends of analogical extension, and current variation indicates that the processes are ongoing. Although the trend with the French vocabulary as a whole has been in the direction of extending the range of the preconsonantal variants, that is, the vowel-final forms, contrary trends can be observed here, perhaps, due to the frequency of the phrase-final forms in counting.

[5] Other prefinal consonants had already been lost: coda /l/ was 'vocalized' (to [w]) and merged with the preceding vowel (VALET > *valt > *vaut* > [vo(t)] 'is worth'); preconsonantal /s/ was lost, at first lengthening the preceding vowel (GUSTO > *goust* > *goût* 'taste'). A final nasal was lost after /r/ even before the orthography was fixed: CARNE > *chair* 'flesh', VERME > *ver* 'worm', FURNO > *four* 'oven'.

(25) French numeral allomorphy

		_#C	_#V	_#‖	Observations on redistribution
1	*un*	ɛ̃	ɛ̃.n	ɛ̃	
2	*deux*	dø	dø.z	**dø**	The _#C form is extended to the _#‖ context, replacing *[døs]
3	*trois*	tʀwa	tʀwa.z	**tʀwa**	The _#C form is extended to the _#‖ context, replacing *[tʀwas]
5	*cinq*	sɛ̃	sɛ̃.k	**sɛ̃k**	The _#V form is extended to the _#‖ context, replacing [sɛ̃], cf. *tronc* 'trunk' [tʀɔ̃]. Some speakers extend [sɛ̃k] to the _#C context also.
6	*six*	si	si.z	sis	
7	*sept*	**sɛt**	sɛ.t	sɛt	The _#‖ form is extended to the _#C context, replacing *[sɛ]
8	*huit*	ɥi	ɥi.t	ɥit	
9	*neuf*	**nœf**	nœ.**f**	nœf	*Neuf ans* '9 years' and *neuf heures* '9 o'clock' are pronounced with [nœ.v], doubtless original.[6] Before other words, vowels or consonants, the _#‖ form has been extended.
10	*dix*	di	di.z	dis	The _#‖ form is extended to the preconsonantal context in *dix-neuf* '19' [diznœf], with voice assimilation, and variably also in *dix-sept* '17' [di(s)sɛt]
20	*vingt*	vɛ̃	vɛ̃.t	**vɛ̃**	The _#C form is extended to the _#‖ context, replacing [vɛ̃t]. However, [vɛ̃t] is used preconsonantally before other numerals: *ving*[t]-*deux* '22', *ving*[t]-*cinq* '25', *ving*[t]-*neuf* '29'.
100	*cent*	sɑ̃	sɑ̃.t	**sɑ̃**	As with *vingt*, the _#C form is extended to the _#‖ context, replacing *[sɑ̃t].

Apart from the numerals *six*, *huit*, and *dix*, there are two other words that retain a prepausal form distinct from the preconsonantal form: *plus* 'more' [ply] _#C, [ply.z] _#V, and [plys] _#‖; and *tous* 'all.MASC.PL' [tu] _#C, [tu.z] _#V, and [tus] _#‖.[7] In addition to the general patterns of liaison, to be considered shortly, a handful of other words retain allomorphs, now suppletive, that originated in sandhi alternations. *Œuf* 'egg' [œf] loses its [f] in the plural *œufs* [ø]; likewise, *bœuf* 'ox' [bœf] – *bœufs* [bø], *os* 'bone.SG' [ɔs], *os* 'bone.PL' [o]. Here the original preconsonantal allomorph occurred before plural /+s/, itself subsequently lost except in liaison environments; now [œf], [bœf], and [ɔs] are used word-finally before a consonant: *bœuf braisé* 'braised steak', *os brisé* 'broken bone'. *Beau* 'fine.MASC' [bo] has a prevocalic suppletive allomorph *bel* [bɛl], as in *un bel été* 'a beautiful summer', plural *beaux* [bo] ~ [bo.z]; and *vieux* 'old.MASC' [vjø] has a prevocalic allomorph *vieil* [vjɛ.j], as in *un vieil arbre* 'an old tree', plural *vieux* [vjø] ~ [vjø.z].

The conditions for the occurrence of liaison are outlined by Tranel (1987: 171) in the following words:

[6] See fn. 4.

[7] Many of the words that retain (prevocalic) liaison alternants belong to categories that are syntactically barred from phrase-final position, such as, determiners, prenominal quantifiers, and prepositions.

> "The phonetic appearance of linking consonants is subject to various conditions that can be divided into four groups of factors: phonetic, syntactic, morphological, and stylistic. There is really only one absolute constraint concerning the appearance of linking consonants: it is of a phonetic nature. Liaison may occur only before a vowel-initial or a glide-initial word. Apart from this, liaison is an extremely variable phenomenon where stylistic factors combine with other factors to yield a considerable range of possibilities going from an extremely limited liaison system to a very dense one. As a rule, the more elevated the style, the more often liaison occurs; the more colloquial the style, the less often liaison occurs. Liaison also depends on the syntactic cohesion between words; the tighter the syntactic link between contiguous words, the more likely liaison is to occur; the looser the syntactic link between contiguous words, the less likely liaison is to occur. Finally, liaison tends to occur more readily if it signals a precise morphological mark (for example, the plural) than if it represents no particular grammatical information."

In contemporary standard French, liaison is almost entirely restricted to the consonants [n], [t] and [z], of which [t] and [z] often realize inflectional morphemes. Otherwise, there are just three words that may link final /ʀ/, two that may link final /t/, and one that may link final /g/ (26a-c) (Tranel 1987: 174-5).

(26) a. *léger* 'light': *un léger incident* 'a slight mishap' [ẽ leʒe.ʀ ẽsidɑ̃]
premier 'first': *au premier étage* 'on the first floor' [o pʀœmje.ʀ etaʒ]
dernier 'last': *un dernier avertissement* 'a final warning' [ẽ dɛʀnje.ʀ avɛʀtismɑ̃]

b. *trop* 'too (much)': *trop aimé* 'too much loved' [tʀo.p eme]
beaucoup 'much': *beaucoup aimé* 'much loved' [boku.p eme]

c. *long* 'long': *un long été* 'a long summer' [ẽ lɔ̃.g ete]

In fact, in the case of *long* (26c), an archaic variant [lɔ̃k] – [ẽ lɔ̃.k ete] – reflects the expected neutralization of voice (with stops realized voiceless) in word-final position. If it were not for the orthography, one might rather say that the feminine form of the adjective, *longue* [lɔ̃g], has taken the place of the masculine form in prenominal position (Morin 1986: 199). This distribution would thus parallel that found with *beau* 'fine' and *vieux* 'old' mentioned above, where the masculine prevocalic form is, in modern standard French, phonologically identical with the feminine form of the adjective.[8]

The syntactic contexts in which obligatory or preferred liaison takes place in standard French can be summarized as follows (after Tranel 1987: 189, Burov 2012: 155-158). The feature in common is close syntactic link with frequent collocation; but not all close syntactic links or frequent collocations demand liaison,[9] and liaison may be acceptable or preferred in some contexts not mentioned here.

(27) i) In a noun phrase with a lexical noun, liaison occurs in the element or elements before it.
ii) In a verb phrase, liaison occurs among the verb and the pronominal satellites around it, and among the pronouns themselves.
iii) Liaison occurs with (most) monosyllabic prepositions, adverbs, and auxiliaries (*avoir*/*être*) and their complements.
iv) Liaison is standard in several fixed collocations.

[8] *Bel* = *belle* [bɛl] , *vieil* = *vieille* [vjɛj]; thus *long* = *longue* [lɔ̃g].

[9] For example, *quand* 'when' (relative adverb) takes liaison —*quand il travaille* 'when he is working' [kɑ̃.t il tʀavaj], but not *quand* 'when' (interrogative adverb), unless followed by *est-ce que*: *quand êtes-vous né?* 'when were you born? [kɑ̃ ɛt vu ne], but *quand est-ce que vous êtes né* 'when were you born? [kɑ̃.t ɛskœ vu.z ɛt ne].

(28) Examples of (27i): *un enfant* ‘a child’ [ɛ̃.n ɑ̃fɑ̃], *aux étudiants* ‘to the students’ [o.z etydjɑ̃], *les autres enfants* ‘the other children’ [le.z otʀœ.z ɑ̃fɑ̃], *mes anciens étudiants* ‘my old students’ [me.z ɑ̃sjɛ̃.z etydjɑ̃]
ces instants ‘these moments’ [se.z ɛ̃stɑ̃], *quelques instants* ‘a few moments’ [kɛlkœ.z ɛ̃stɑ̃], *quelles affaires?* ‘what matters?’ [kɛl.z afɛʀ]
un gros arbre ‘a thick tree’ [ɛ̃ gʀoz. aʀbʀ], *de vieux amis* ‘some old friends’ [dœ vjø.z ami]

(29) Examples of (27ii): *ils arrivent* ‘they are arriving’ [i.z aʀiv], *prends-en* ‘take some’ [pʀɑ̃.z ɑ̃], *nous en avons* ‘we have some’ [nu.z ɑ̃.n avɔ̃], *dont on a parlé* ‘of which we spoke’ [dɔ̃.t ɔ̃.n a paʀle], *vient-il?* ‘is he coming?’ [vjɛ̃.t il]

(30) Examples of (27iii): *dans un mois* ‘in a month’ [dɑ̃.z ɛ̃ mwa], *en anglais* ‘in English’ [ɑ̃.n ɑ̃glɛ], *en écoutant* ‘(while) listening’ [ɑ̃.n ekutɑ̃], *chez elle* ‘at her house’ [ʃe.z ɛl], *très intéressant* ‘very interesting’ [tʀɛ.z ɛ̃teresɑ̃], *moins autoritaire* ‘less authoritarian’ [mwɑ̃.z ɔtɔʀitɛʀ], *c’est impossible* ‘it is impossible’ [sɛ.t ɛ̃pɔsibl], *j’y suis allé* ‘I went there’ [ʒi sɥi.z ale].

(31) Examples of (27iv): *accent aigu* ‘acute accent’ [aksɑ̃.t egy], *nuit et jour* ‘night and day’ [nɥi.t e ʒuʀ], *États-Unis* ‘United States’ [eta.z yni], *tout à coup* ‘all of a sudden’ [tu.t a ku], *de haut en bas* ‘from top to bottom’ [dœ o.t ɑ̃ bɑ], *de temps en temps* ‘from time to time’ [dœ tɑ̃.z ɑ̃ tɑ̃].

If the sketch of historical development given above is correct, in the French of c. 1500, all consonant-final words will have had at least two variants —one before vowel initial words, retaining the final consonant, and one before consonant-initial words lacking it. In later French, aside from the cases where liaison retains a prevocalic allomorph, one or other of the alternatives has been extended to all contexts. More often than not such an invariant form is the one with deletion of the original (Old French) final consonant, but in many particular cases the consonant is retained. A few contrasting pairs are given in (32). The unpredictability is greatest among nouns; verbs and adjectives nearly always lost a final consonant. In words borrowed or formed in French after the Renaissance, an orthographic final consonant is generally pronounced.

(32)

broc ‘pitcher’ [bʀo]	*bloc* ‘block’ [blɔk]
clerc ‘clerk’ [klɛʀ]	*turc* ‘Turk, Turkish’ [tuʀk]
clef ‘key’ [kle]	*nef* ‘nave’ [nɛf]
nerf ‘nerve’ [nɛʀ]	*serf* ‘serf’ [sɛʀf] (or [sɛʀ])
encens ‘incense’ [ɑ̃sɑ̃]	*sens* ‘meaning’ [sɑ̃s]
tas ‘pile’ [tɑ]	*as* ‘ace’ [ɑs]
avis ‘opinion’ [avi]	*vis* ‘screw’ [vis]
chaos ‘chaos’ [kao]	*rhinocéros* ‘rhinoceros’ [ʀinɔseʀɔs]
sot ‘stupid’ [so]	*dot* ‘dowry’ [dɔt]
chat ‘cat’ [ʃa]	*mat* ‘matt’ [mat]
goût ‘taste’ [gu]	*août* ‘August’ [ut] (or [u])
début ‘beginning’ [deby]	*brut* ‘rough’ [bʀyt] *but* ‘goal’ [byt] (or [by])

3.3.2 Initial geminates from coda assimilation: *rafforzamento fonosintattico*

A striking morphophonological characteristic of Central and Southern Italian varieties (including standard Italian) and of Sardinian is the gemination, or strengthening, of initial consonants after specified vowel-final words (*rafforzamento*, or *raddoppiamento*, (*fono*)*sintattico*, abbreviated here RF; Marotta 2010), e.g. It. *da casa* ‘from home’ [da

k.ˈkasa] vs. *la casa* ‘the house’[la ˈkasa].[10] The origin of this phenomenon is clear enough (Loporcaro 1997): the set of vowel-final words that trigger RF incudes (all?) those that, in Latin, (i) ended in a consonant, and (ii) were monosyllabic. A few of these are unstressed, and proclitic (33a). Compounds whose final part is one of these also trigger RF (33b).

(33) a. *a* < AD ‘to’: It. *a lui* ‘to him’ [a l.ˈluj]; Neap. *a Napule* ‘to Naples’ [a n.ˈnapule][11]
che < QUID × QUOD ‘that’: *dice che vuole* ‘she says she wants to’ [ˈditʃe ke v.ˈvwɔle]; *i bambini che vedi* ‘the children that you see’ [i bamˈbini che v.ˈvedi]
da < DE+AB, DE+AD: *da me* ‘from me’ [da m.ˈme], *da fare* ‘to do’ [da f.ˈfare]
e < ET ‘and’: *eppure* < *e* + *pure* ‘and yet’ [ep.ˈpure]
né < NEC ‘neither/nor’: *né caldo né freddo* ‘neither hot nor cold’ [ne k.ˈkaldo ne f.ˈfreddo]
o < AUT ‘either/or’: *ossia* < *o* + *sia* ‘that is to say’ [os.ˈsia]
se < post-Imperial Latin SED < SI × QUID (Loporcaro 197: 27) : *se puoi* ‘if you can’ [se p.ˈpwɔj]

b. *come* < QUOMODO+ET ‘like, as’: *come noi* ‘like us’ [ˌkome n.ˈnoj]
qualche < QUALE+QUID ‘some’: *qualche giorno* ‘some day’ [ˌkwalke d.ˈʤorno]; Neap. *quacche libbro* ‘some book’ [ˈkwakkə l.ˈlibbrə]

Others of the original triggers are stressed words (34). With a similar effect are some original consonant-final words that became monosyllabic by aphaeresis within Romance: *là* < (IL)LAC ‘there’, *lì* < (IL)LIC ‘there’, *qua* < (EC)CUM HAC ‘here’, *qui* < (EC)CUM HIC ‘here’, *ciò* < ECCE HOC ‘this’, e.g. Neap. *llà bascio* ‘down there’ [ˌlla. b.ˈbaʃʃə]

(34) *che* < QUID: *che dite?* ‘what are you saying?’ [ˌke d.ˈdite]
dà < DAT: *dà molto* ‘gives a lot’ [ˈda m.ˈmolto]
dì < DIC: *dimmi* ‘tell me’
è < EST: *è vero* ‘it is true’ [ˌɛ v.ˈvero]
fa’ < FAC: *fammi un favore* ‘do me a favour’
più < PLUS: *più morto che vivo* ‘more dead than alive’ Florence [ˌpju m.ˈmɔrto he v.ˈvivo]; Neap. *cchiù doppo* ‘later’ [kˌkju d.ˈdoppə]
sì < SIC; *così* < (EC)CUM SIC ‘so’: *così male* ‘so bad’ [koˌsi m.ˈmale], *cosiddetto* ‘so-called’
sta < STAT: *sta bene* ‘it is good’ [ˌsta b.ˈbɛne]
tre < TRES: *tre cani* ‘three dogs’ [ˌtre k.ˈkani]

At an early stage, a Latin final consonant in such contexts became totally assimilated to a following onset consonant. In Classical Latin this process was represented orthographically in the items here only with AD as a prepositional prefix: AD + FIGERE → AFFIGERE ‘to attach’ AD + SUMERE → ASSUMERE ‘to take on’. But inscriptional evidence from the Imperial period confirms that the process was more widespread (Loporcaro 1997: 42, 121): <*at tuos*> for AD TUOS, <*sud die*> for SUB DIE. In proto-Italo-Romance, in Loporcaro’s interpretation (1997: 121) there will have been paradigms such as (35).

[10] Consonants that are inherently geminate in Italian: /ɲ/ [ɲ.ɲ], /ʎ/ [ʎ.ʎ], /ʃ/ [ʃ.ʃ], /ʦ/ [t.ʦ], /ʤ/ [d.ʤ] are unaffected, as is /s/ in an initial cluster /sC-/. Gemination by RF is noted orthographically only in cases of ‘univerbation’, where two originally separate words are conventionally written as one, as with *eppure*, *ossia*, in (33a). Though Italian linguists generally transcribe the long or geminate consonants with [ː] we prefer to write [C.C] which reveals the syllable structure. We do not mark vowel length: by default, a vowel in an open syllable before a consonant onset is long, otherwise short.

[11] Neapolitan examples are from Ledgeway (2009: 40-48).

(35)	a. _#C	b. _#V	c. _#‖
	/ˈdat#ˈpane/ ‘gives bread’	/ˈdat#ˈakwa/ ‘gives water’	/ˈnon#ˈdat/ ‘does not give’
	[ˈda p.ˈpane]	[ˈda(t) ˈakwa]	[ˈnon ˈda(t)]

The evidence suggests that Latin word-final consonants were retained long enough in monosyllables, or the few stressed finals in polysyllables, to trigger RF effects, while word-final consonants elsewhere had already been deleted in Italo-Romance. The special status of monosyllables in this respect is consistent with what is observed more widely in Romance with Latin final <m> or <n>, which are often preserved in Romance monosyllables as /n/, and lost elsewhere; e.g. It. *non* < NON ‘not’, *son*(*o*) < SUM ‘I am’, Sp. *quien*, Rom. *cine* < QUEM ‘who?’, Fr., Cat. *ton* < *tum < TUUM ‘your.2SG’, Oc. *ren*, Fr. *rien* < REM ‘nothing’, Sp. *tan* < TAM ‘so’; but It. *nome*, Cat., Oc. *nom* < NOMEN ‘name’, It. *sciame*, Cat. *eixam* < EXAMEN ‘swarm’, and no sign of Latin final <-*m*> in polysyllables.[12] When word-final consonants are deleted in contexts such as (35b, 35c), the geminate in contexts such as (35a) remains as a relic. It is no longer recoverable why *dà* ‘gives’ and the other words of (33) and (34) trigger RF, but, for example, *di* ‘of’ < DE, *mi*/*me* < ME ‘me’, *ti*/*te* < TE ‘thou’, *si*/*se* < SE ‘3.REFL.’, *la* ‘DEF.FEM.SG’ do not. (However, *e* ‘and’ retains a prevocalic variant *ed* in some contexts, e.g. *titoli ed esami* ‘titles and examinations’, and *a* ‘to’ a variant *ad*: *passare ad altro* ‘change the topic’.)

In Italo-Romance, the present tense paradigms of *dare* ‘give’, *stare* ‘stand’, *fare* ‘do’, *avere* ‘have’, *sapere* ‘know’, and *andare* ‘go’, influence one another in analogical reformations, giving rise to 3SG monosyllabic forms *fa* (as if from *FAT, for FACIT), *ha* (as if from *HAT, for HABET), *sa* (as if from *SAT, for SAPIT), and *va* (as if from *VAT, for VADIT), triggering RF just like the models *dà* (< DAT) and *sta* (< STAT). Likewise, the 2SG imperatives *da*, *sta*, and *va* come to trigger RF on the model of *fa’* (< FAC) and *dì* (< DIC). The fact that *ha* is one of the early RF triggers has an important morphological consequence through its role in forming the future tense, as in *farà* ‘will do.3SG’ (*fare* + *ha*), *andrà* ‘will go.3SG’ (*andare* + *ha*), *finirà* ‘will finish.3SG’ (*finire* + *ha*). In our interpretation of the historical sequence of events, the 3SG future form serves to expand the set of RF triggers beyond monosyllables to include polysyllables with stressed final vowels. Hence in standard Italian (36):

(36) *andrò piano* ‘I will go slowly’ [anˈdrɔ p.ˈpjano], *caffè forte* ‘strong coffee’ [kafˈfɛ f.ˈfɔrte], *città bella* ‘beautiful city’ [ʧitˈta b.ˈbɛlla]

In standard Italian, a further consequence is an almost complete remotivation of RF on phonological grounds: in this variety, RF is triggered by (a) all words with a stressed final vowel (including monosyllables: *sto male* ‘I am not well’ [ˌstɔ m.ˈmale]), and (b) the small original list of items (33), to which a handful of others with an unstressed final vowel are added, apparently through analogies which are not wholly clear: *ma* ‘but’ (< MAGIS), *tra* < INTRA and *fra*[13] < INFRA ‘between, within’, *sopra* ‘above’ < SUPRA, *contra* ‘against’ < CONTRA, and *dove* ‘where’ < DE + UBI. Other Italo-Romance varieties lack this remotivation; in them, RF is triggered almost exclusively by words that had a final consonant in Latin. But since the lexical division is opaque, the list of trigger words can vary markedly between dialects, as items are analogically added to or removed from the list (Loporcaro 1997: 72-117). Or the remotivation is itself undermined by sound change: in Tuscan, word-final

[12] However, IAM ‘already’ loses its nasal in all Romance varieties: Sp. *ya*, Fr (dé)*jà*, so Loporcaro may not be correct to include It. *già* among ‘i monosillabi oggi raddoppianti riconducibili ad etimo con finale consonantica’ (1997: 49).

[13] Loporcaro (1997: 24) includes *tra* and *fra* among words having a Latin final consonant, though he does not offer etyma, and they are usually traced to INTRA and INFRA.

stressed diphthongs lose their glide element, but fail to trigger RF as other final stressed monophthongs do, giving e.g. Florence *vorrei parlare* 'I would speak' [vorˈrɛ ɸarˈlare] alongside *vorrà parlare* 'she would speak' [vor.ˈra p.parˈlare] (Loporcaro 1997: 11).

The status of RF is somewhat different in Sardinian, where, alongside vowel-final RF triggers such as *a* 'to', *e* 'and', *ne* 'neither', corresponding to those in Italian, 3SG verb forms retain allomorphs with reflexes of Latin <-*t*>. Hence we may find alternations such as these (37a) in Logudorese (Contini 1986: 531, from Nughedu, except the first example from Berchidda). And reflexes of plural <-*s*> are retained, where intervocalic [z] alternates with RF (37b). Note the grammaticalization of the contrast between lenition and RF to mark nominal number in (37c).[14]

(37)	a.	e b.ˈbenniðu 'he came'	ˈbenniðu ˈɛstɛ? 'did he come?'
		fi p.pasˈtɔrɛ 'he was a shepherd'	pasˈtɔrɛ ˈviði? 'was he a shepherd?'
		a s.siˈɣið a f.faeˌ̯ˈɖarɛ 'he continued speaking'	siˈɣiðu ˈaða? 'did he continue?'
		ˈfaɣɛ t.ˈtempu m.ˈmalu 'the weather is bad'	ˈtempu m.ˈmalu ˈvaɣɛðɛ 'is the weather bad?'
	b.	ˈdua f.ˈfeminaza 'two women'	ˈsun ˈduaza 'they are two.FEM'
		ˈbattɔ k.ˈkanɛzɛ 'four dogs'	ˈsun ˈbattɔrɔ
		sɔˈz annɔzɔ 'the years'	
	c.	sa t.ˈtanka d.dɛ anˈtoni 'Antoni's properties'	sa ˈðanka ðɛ anˈtoni 'Antoni's property'

Like French liaison, Italo-Romance *rafforzamento fonosintattico* illustrates how quite straightforward, phonologically well motivated, sandhi phenomena can become opaque, and then suffer analogical lexical readjustments. In this respect it is illustrative to analyse the current changes underway in Andalusian Spanish due to the historical aspiration of /s/ in the coda of the syllable. Currently this aspiration is mostly not present in C#‖contexts (*más* [ma] 'more') given its low perceptual salience and potential to be confused with breathy voice which usually occurs at the end of utterances (O'Neill 2005); it appears more consistently in C#.V contexts depending on the variety (*mas agua* [ma.ˡha.ɣwa] 'more water'). In C#.C contexts it can modify the manner of articulation of the following consonants and produce novel phonemic distinctions (O'Neill 2010). The developments are generally that in eastern varieties all following consonants with the exception of fricatives are geminated and voiceless occlusives are aspirated also. In western varieties the same holds true but the elongation of the consonants is much less (for the aspirated voiceless stops it is only variably present) and the consonants can be pronounced with breathy voice (nasals) or frication (laterals and spirants). Whilst these changes are attested both word internally and across words (*caco* [ˡka.ko] 'thief', *casco* [ˡka.kʰo] 'helmet' *dos comiendo* [ˡdo.kʰo.ˡmjen̪.d̪o] 'two people eating') some interesting developments are occurring in the city of Seville. Certain lexical items show consistent phonetic cues of aspiration for orthographic <s> across words (e.g. *más* 'more' and *dos* 'two') but such cues are only variably present for the plural marker (*perro* 'dog' *vs. perros* 'dogs') and totally absent for the 2sg marker in the verb (*come* 'he eats' vs. *comes* 'you eat' see O'Neill (ms). The phonological effects appear to be becoming

[14] Note paragoge in phrase-final position in these examples, alternating with RF of a following initial consonant. Thus *cantat* 'sing.3SG' has three sandhi variants: [ˈkanta.ð] before a vowel, [ˈkanta +RF] before a consonant, and [ˈkantaða] before pause (Loporcaro 1997: 114).

opaque and lexicalised (*comiendo* [ˈko.ˈmjen̪.d̪o] ‘eating’, *dos comiendo* [ˈdo.kʰo.ˈmjen̪.d̪o] ‘two people eating’, *¿vienes comiendo?* [ˈbje.ne.ko.ˈmjen̪.d̪o] *‘*Are you eating now?’, *niños comiendo,* mainly [ˈni.ɲo. ko.ˈmjen̪.d̪o] but also [ˈni.ɲo. kʰo.ˈmjen̪.d̪o] ‘children eating’. Moreover, there is some evidence (Pons-Rodriguez ms) that on the verb post-aspiration is being morphologized (*¿Qué comes?* ‘what are you eating?’ [ˈke.ˈkʰo.me], *¿Qué come?* ‘What is she eating?’ [ˈke.ko.me] which must have its origins in analogy with nouns in which the phonological sandhi effects could also be becoming morphologized (*el caco* ‘the thief’[el.ˈka.ko], *los cacos* ‘the thieves’ [lo.ˈkʰa.ko]. The evidence for Andalusian Spanish is only preliminary and variable but there are pararels with the French liaison and especially RP; the textual and lexical incidence of these phenomena remains very pervasive. In addition, in the Italian case, one might well say RF is linguistically superfluous. Speakers of standard Italian from those areas (Northern Italy) whose local varieties lack RF largely ignore it, without problems of comprehension, or of social stigma.

References

Andersen, H. (ed.) (1986). *Sandhi phenomena in the languages of Europe.* Berlin: Mouton de Gruyter.

Bibiloni, G. (2016). *El català de Mallorca. La fonètica.* Palma: Lleonard Muntaner.

Bonet, E. & Lloret, M.-R. (1998). *Fonologia catalana* (Barcelona: Ariel).

Burov, I. (2012). *Les phénomènes de Sandhi dans l'espace gallo-roman*. Thèse de doctorat Université Michel de Montaigne - Bordeaux III, 2012. Français. <NNT : 2012BOR30041>. <tel-00807535> https://tel.archives-ouvertes.fr/tel-00807535

Contini, M. (1986). 'Les phénomènes de sandhi dans le domaine sarde'. In Andersen (1986: 519-550).

Dols Salas, Nicolau A. (1993). The predictive formalization of consonantal contacts in Majorcan Catalan (empirical and theoretical bases). M.Phil. thesis, Department of Hispanic Studies, University of Sheffield.

Hualde, Ignacio (2013) Intervocalic lenition and word-boundary effects. *Diachronica* 30(2):232-266

Encrevé, Pierre (1988) *La liaison avec et sans enchaînement: Phonologie tridimensionnelle et usage du français*. Paris: Le Seuil.

Ledgeway, A. (2009). *Grammatica diacronica del napoletano.* Tübingen: Max Niemeyer.

Lloret, M.-R. & Jiménez, J. (2007). Prominence-driven epenthesis: evidence from Catalan. ms. https://www.academia.edu/5683215/

Loporcaro, M. (1997). *L'origine del radoppiamento fonosintattico. Saggio di fonologia diacronica romanza.* Basel/Tübingen: Francke.

Marotta, G. (2010) 'Raddoppiamento sintattico'. In Simone, R., Berruto, G. & D'Achille, P. (eds.) *Enciclopedia dell'italiano*. http://www.treccani.it/enciclopedia/raddoppiamento-sintattico_%28Enciclopedia-dell%27Italiano%29/

Morin, Y.-C. (1986). 'On the morphologization of word-final consonant deletion in French'. In Andersen (1986: 167-210.)

O'Neill, P. (2005). Final /s/ in Western Andalusian Spanish: the phonetic neutralization of a phonological contrast. Language Design, 7:151-166.

O'Neill, P. (2010). Variación y cambio en las consonantes oclusivas del español de Andalucía. Estudios de Fonética Experimental, 19:11-41.

O'Neill, P. (under consideration). /s/ aspiration in Andalusian Spanish – an experimental study of 4 cities.

Penny, R. (1986). 'Sandhi phenomena in Castilian and related dialects'. In Andersen (1986: 489-503).

Pons Rodríguez, L. (ms) *Nuevos plurales en el español de Andalucía: Paradigma, gramática y pragmática.*

Recasens, D. (1993). *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.

Sampson, R. (2010). *Vowel prosthesis in Romance. A diachronic study*. Oxford: Oxford University Press.

Sorianello, P. (2010) 'Gorgia toscana'. In Simone, R., Berruto, G. & D'Achille, P. (eds.) *Enciclopedia dell'italiano*. http://www.treccani.it/enciclopedia/fonetica-sintattica_%28Enciclopedia-dell%27Italiano%29/

Tranel, B. (1987). *The sounds of French. An introduction*. Cambridge: Cambridge University Press.

Wheeler, M. W. (2005). *Phonology of Catalan*. Oxford: Oxford University Press.