

# Grammar-internal and Grammar-external Assimilation

John Harris

University College London

johnh@ling.ucl.ac.uk

## ABSTRACT

Processes traditionally described as assimilations fall into two main types according to the impact they have on the informational content of speech signals. In grammar-internal assimilation, exemplified by vowel harmony, sound properties that are suprasegmentally extended have a linguistic marking function. In grammar-external assimilation, exemplified by lenition, the extended properties belong to the carrier signal.

## 1. SOUND OVERLAP

Where does ‘phonological’ assimilation stop and ‘phonetic’ coarticulation start? It is clear that the two notions are closely related: both refer to some kind of overlap between neighbouring sounds, and any process described as assimilatory is inevitably accompanied by coarticulatory effects. The supposed distinction can be interpreted in various ways. For example, assimilation might be viewed as belonging in the grammar, while coarticulation belongs outside it. Or assimilation might be deemed to operate deep in the grammar (at the ‘lexical’ level, say), while coarticulation operates at the periphery (‘postlexically’).

A standard listing of criteria for classifying individual cases of overlap in terms of this distinction might run as follows. Coarticulation is phonetically continuous, does not neutralise phonological contrasts, and is lexically exceptionless. Assimilation is phonetically discrete, neutralises contrasts, and may have exceptions.[1]

However, this idealised classification is difficult to sustain when individual cases of overlap are scrutinised.[2] A consistent cut-off point, while retaining a certain intuitive appeal, remains elusive, leading some to conclude that assimilation and coarticulation are no more than informal labels for a unitary phenomenon. In articulatory phonology, for example, dynamically specified gestures can overlap to varying extents, thereby producing sound sequences with varying degrees of perceptual overlap.[3] In phonetically driven constraint-based theory, categorical or neutralising effects of overlap emerge as a result of ranked or weighted constraints operating on concrete, phonetically continuous parameters.[4][5]

While the concrete parameters, such as VOT or position in the vowel space, can be measured in terms of continuously changing values, their behaviour in speech perception and in phonological systems is anything but continuous. A good example is 20ms boundary in the perception of VOT contrasts. Moreover, the number of perceptual and phonological categories evinced by the discontinuities on each of these parameters is comparatively small. Take the case of vowel height, particularly relevant here in that it is

implicated in one widespread type of assimilation, namely height harmony. Vowel height can be synthesised in terms of continuously varying formant frequencies, and listeners can detect quite minute differences along this parameter. But the kind of information carried by such fine differences appears to be non-linguistic, providing for example indexical or personal marking (as demonstrated for example by Labov’s work on the raising of English short *a* in northern US cities).[6] Where quality differences do bear a linguistic marking function, categorisation is much coarser. Thus phonological systems are limited to three vowel-height categories, with the possibility of a subsidiary tense-lax distinction. At some point, a model using phonetically continuous features has to stipulate a limit on how fine-grained the distinctions along a particular parameter can be. As far as the parameters involved in assimilation are concerned, this essentially leads us back to the problem of determining where phonetic gradience stops and categorical behaviour starts.

## 2. PHONOLOGICAL KNOWLEDGE

Suppose we use instead the linguistic marking function just mentioned as a criterion for deciding whether an individual case of sound overlap counts as a grammar-internal assimilation. Rather than asking whether the overlap is phonologically neutralising and/or phonetically discrete, we ask whether it makes any contribution to the specifically linguistic content of speech signals.

This raises a wider question, regarding the knowledge contained in a listener-talker’s phonological grammar. In particular, does the grammar contain both of the following?

- Phonological knowledge (1)
- (a) Conventionalised knowledge that enables the listener-talker to extract linguistic information from speech signals.
  - (b) Knowledge of the phonetically natural pressures that shape the sound material represented under (1)a.

Any model of phonological grammar must minimally contain (1)a. The view that phonological knowledge extends to (1)b is reflected in proposals that the grammar can contain such devices as markedness conditions and phonetically driven constraints. These devices recapitulate explanations of sound patterns that are provided by research in the realms of speech production, auditory perception, historical change, and so forth.

Phonetically natural pressures act on the systems employed in the transmission and reception of linguistic information, but in and of themselves they have no linguistic-informational value. This point informs an alternative,

‘minimalist’ view of phonological knowledge, one in which the grammar exclusively contains (1)a.

Narrowing the focus to (1)a allows for a consistent classification of assimilatory processes that differs quite radically from anything associated with standard descriptions. Some cases, represented below by vowel harmony, must be allocated to (1)a on the grounds that the sound properties given suprasegmental scope by assimilation are linguistically informative. Others must be excluded from (1)a for the reason that they involve the temporal extension of properties that lack a linguistic marking function. Under the view that phonological knowledge is restricted to (1)a, the distinction can be characterised simply as grammar-internal versus grammar-external. Under the view that phonological knowledge is more encompassing than this, a corresponding distinction can only be maintained if knowledge of type (1)a is assumed to occupy some independent sub-module within the grammar.[7]

The specific examples to be discussed below both qualify as assimilation in much of the relevant literature. They are (i) a particular version of the phenomenon of height harmony already touched and (ii) a type of lenition whereby an intervocalic obstruent assimilates to the voicing and open structure of the surrounding vowels.

### 3. INFORMATION IN THE SPEECH SIGNAL

What does it mean to say that a particular process has a linguistic marking effect? The question can be answered by determining how the process impacts on the informational content of speech signals.

Speech can be seen as linguistically informative modulations of a carrier signal.[8] The carrier signal, primarily reflective of personal and expressive qualities, is associated with a neutral vocal-tract shape, is typically (though not necessarily) voiced, and is neutral with respect to linguistic information. Linguistically significant modulations of this carrier signal have informative value due to the principle that change is perceptually more salient than stability.[9] The ease with which a modulation can be detected is dependent on its relative magnitude, as measured in terms of its trajectory though an acoustic space defined by the parameters of spectral shape, amplitude, periodicity, and fundamental frequency.

Any phenomenon traditionally described as assimilation can in principle take one of two forms, depending on whether the sound property that is temporally extended by assimilation belongs to the linguistic content of a speech signal or to the carrier signal. Under the view that phonological knowledge is exclusively of the kind described in (1)a, only the first of these counts a grammar-internal assimilation. Vowel harmony, as we will now see, falls into this category. Lenition does not.

### 4. VOWEL HEIGHT HARMONY

In Sesotho (southern Bantu), mid lax (non-ATR) vowels become tense (ATR) under the influence of a high vowel in a following syllable. This is illustrated in (2)a, where stem vowels can be seen to raise under the influence of a high

suffix vowel. As shown in (2)b, the effect is unbounded: any continuous span of mid-vowelled syllables undergoes raising when a high-vowelled trigger is present.

(a)	INFINITIVE	CAUSATIVE	(2)
	<i>reka</i>	<i>rekisa</i>	‘buy’
	<i>pota</i>	<i>potisa</i>	‘go round’
(b)	INFINITIVE	NEGATIVE	
	<i>xebetlela</i>	<i>xebetleli</i>	‘be fed’
	<i>kokota</i>	<i>kokoti</i>	‘knock’

This pattern clearly involves anticipatory assimilation and/or coarticulation: an adjustment to the vowel-space position for mid vowels shortens the trajectory to the position required for a following high vowel.

As to the question of whether the raised quality of mid vowels in Sesotho has a linguistic marking function, the answer must be yes. The qualitative change in stems affected by raising harmony provides a contrast that helps cue the grammatical category of the word in which the stem is located. The spectral pattern associated with raised quality thus counts as a linguistically significant event in the speech signal.

The linguistic functionality of raised quality is further confirmed by the fact that it is retained even when the assimilatory trigger is absent. This effect results from the syncope of *i* in certain coronal contexts in Sesotho. The examples in (3) contain the causative suffix already illustrated in (2).

	INFINITIVE	CAUSATIVE	(3)
	<i>tlola</i>	<i>tlotsa</i> (< * <i>tlodisa</i> )	‘smear oneself’
	<i>laela</i>	<i>laetsa</i> (< * <i>laedisa</i> )	‘order’

The assimilation itself, it is reasonable to conclude, is also linguistically significant and must therefore be represented in the phonological grammar. The suprasegmental scope of the linguistic information borne by harmonically raised quality serves to demarcate the domain over which the morphological category associated with a given affix operates. The specific means by which this long-distance effect is captured is not immediately relevant, be it by feature spreading, multiple feature copying, or the non-procedural deployment of suprasegmental features.

Given their grammatical marking function, the alternations caused by vowel raising must be part of the conventionalised knowledge a Sesotho listener-talker draws on to parse the linguistic content of speech signals. It is a quite independent matter whether additional cognitive provision needs to be made for the phonetic naturalness of the process, for example in the form of some constraint penalising the expenditure of articulatory effort.[5] In a model of grammar containing only knowledge of type (1)a, there is no place for extra information of this sort. This is not because the naturalness of the process is compromised by the opacity caused by *i*-syncope (though of course it is) but because it contributes nothing to the interpretation of specifically linguistic information.

Grammar-internal height harmony exploits a sound resource that no doubt has extra-grammatical origins. The diachronic transition into the grammar presumably involves a process

whereby listener-learners reinterpret the mechanical (that is, non-cognitive) effects of coarticulation as part of the intended message.[10]

## 5. LENITION

Consider the following data from three genetically unrelated languages (accents mark tone in the Basaa and Ibibio examples;  $\beta$  and  $\gamma$  stand for frictionless continuants):

- |                          |           |              |                  |
|--------------------------|-----------|--------------|------------------|
| (a) English              |           |              | (4)              |
| fit                      | fi[t]     | fitter       | fi[r]er          |
| get                      | ge[t]     | get on       | ge[r] on         |
| (b) Basaa (Narrow Bantu) |           |              |                  |
| <i>kap</i>               | ‘share’   | <i>keβa</i>  | ‘be shared’      |
| <i>tet</i>               | ‘grind’   | <i>tira</i>  | ‘be ground’      |
| <i>lok</i>               | ‘lie’     | <i>luya</i>  | ‘be lied’        |
| (c) Ibibio (Lower Cross) |           |              |                  |
| <i>déép</i>              | ‘scratch’ | <i>dééβé</i> | ‘not scratching’ |
| <i>kóót</i>              | ‘call’    | <i>kóóró</i> | ‘not calling’    |
| <i>fáák</i>              | ‘wedge’   | <i>fááyá</i> | ‘not wedged’     |

These examples illustrate a general type of lenition by which an oral stop is converted into some kind of continuant (spirantisation) that is typically voiced (sonorisation) and may have a relatively open articulation (vocalisation).

The wide distribution of this phenomenon across different languages suggests that it is phonetically natural. This is usually taken to mean that it involves assimilation and/or coarticulation: the vocal-fold vibration and open vocal-tract shape associated with the vowels are imposed to varying degrees on the intervening consonant.[5][11]. On the other hand, that this cannot be a universal, mechanical effect is confirmed by the fact that not all languages exhibit this type of entrenched lenition.

Further, in the languages illustrated here, lenition does not occur freely in all intervocalic contexts (as is sometimes claimed to be the case in Spanish for example). To see this, now consider the additional examples in (5). Here the Basaa and Ibibio forms contain oral stops in stem-initial position following a prefix vowel.

- |               |          |            |     |
|---------------|----------|------------|-----|
| (a) English   |          |            | (5) |
| retain        | re[t]áin | *re[r]áin  |     |
| (b) Basaa     |          |            |     |
| <i>li-pan</i> | *liβan   | ‘forest’   |     |
| <i>li-tám</i> | *liram   | ‘fruit’    |     |
| (c) Ibibio    |          |            |     |
| <i>ú-táŋ</i>  | *úráy    | ‘plaiting’ |     |
| <i>ú-káp</i>  | *úγáp    | ‘covering’ |     |

For English, the difference between (4)a and (5)a shows that tapping is sensitive to metrical foot structure. An intervocalic coronal stop is susceptible to tapping only if occurs outside the prominent (i.e. initial) syllable of a foot. Hence tapping occurs in (4)a (where  $t$  is foot-medial or final) but not in (5)a (where  $t$  is foot-initial).

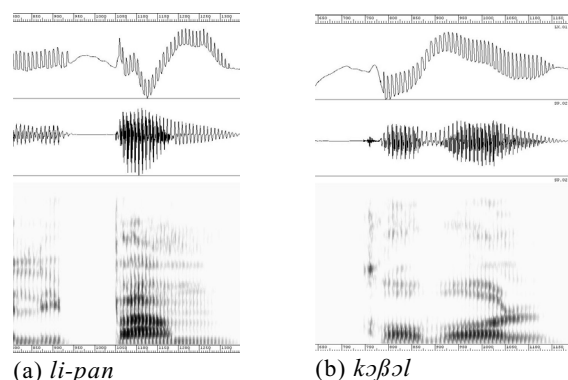
A comparison between (4) and (5) reveals a similar situation in Basaa and Ibibio: intervocalic oral stops resist lenition when initial in the stem. The most obvious difference with respect to English is that, being tone languages, Basaa and

Ibibio lack stress prominence. However, the parallel with English becomes more striking when we take into consideration the fact that the size of the stem in both languages is constrained by a templatic limit that is equivalent in size to a heavy-light trochaic foot.[12]

These cross-linguistic similarities suggest that segmental lenition is one of the properties by which prominence relations are signalled. Furthermore, in exhibiting sensitivity to prosodic and morphological domain structure, lenition in our three illustrative languages must be considered to have linguistic functionality. In Basaa and Ibibio, the distribution of lenited versus unlenited consonants provides information about the location of stem boundaries. In VCV sequences, an unlenited plosive marks the beginning of a stem. In English, an aspirated plosive  $t$  marks the beginning of a foot. The strong affinity between feet and words in English makes this a potential source of information in morphosyntactic parsing. Thus the effects of lenition must be represented in some form or another in the phonological grammars of all three languages.

If the effects of intervocalic lenition have grammatical presence, does that necessarily mean they are represented in the form of assimilation, just as with vowel harmony? The view that lenition does warrant this kind of treatment is widespread enough to have achieved the status of textbook orthodoxy.[13] It is reflected in analyses in which some combination of values for the features [voice], [continuant] and [sonorant] are spread or copied from the flanking vowels onto the target consonant.

However, treating lenition as grammar-internal assimilation overlooks an important difference with respect to vowel harmony. When applied to vowels, the standard feature specifications [+voice], [+continuant], and [+sonorant] identify linguistically insignificant, background properties of the speech signal. That is, they refer to the neutrally open vocal-tract shape associated with the carrier signal – unlike features such as [round] or [low], which correspond to linguistically significant signal modulations. Indeed the periodicity referred to by [+voice] is not even a defining property of vowel quality (think of whispered vowels).



**Figure 1:** Basaa *li-pan* ('forest') vs. *kəβəl* ('hang') (male speaker): Lx (top), audio (middle), broadband spectrogram (bottom).

The effect of lenition on the informational content of the speech signal speaks against this account. Intervocalic

spirantisation, sonorisation, and vocalisation reduce the magnitude of modulations across VCV sequences. In intervocalic position, a lenited consonant blends more closely with the background represented by the carrier signal than does an unlenited congener. This can be seen in Figure 1, where we can compare lenited and unlenited versions of labial consonants in Basaa. An array of acoustic cues makes an intervocalic plosive stand out in stark relief against the carrier signal (see Figure 1a): an abrupt and sustained drop in amplitude, rapid formant transitions in the approach and release phases, the presence of a release burst, the frequency of that burst, etc. Of these cues, only the formant transitions are retained by the corresponding lenited sound (see Figure 1b). As a result, the trajectory across the VCV sequence is that much smoother.

A more plausible take on lenition is thus to view it as involving the loss of linguistic information rather than the extension of non-linguistic properties. The most direct way of representing this effect in the grammar is in terms of the deletion or suppression of features, not in terms of feature spreading or copying. It is not immediately clear how this might be achieved with standard bivalent features, which force us to treat any kind of lenition as the replacement of one set of values by another – thereby implying that a lenited segment projects just as much information as an unlenited congener. An alternative is to capture lenition directly in terms of the loss of components from phonological representations.[12][14] Thus, for example, the labial stop in Basaa *lipan* has several components, while the continuant in *kɔβəl* has only one. While the continuant is not phonologically assimilated to its surrounding segments, it is less distinct from its neighbours than the stop would be.

The occurrence of lenited *β* in Basaa *keβa* ('be shared', cf. *kap* 'share') is part of what signals that this word is passive. But this is not to say that the passive is linguistically signalled by voicing, continuancy, and sonorancy being spread across the morpheme, for the simple reason that these effects are not linguistically signalled in this word at all; they are part of the carrier signal.

While the effects of entrenched lenition must be represented in the phonological grammar, they should not be represented in the form of grammar-internal assimilation. That is not to deny the role of assimilatory or coarticulatory factors in the origination of the effects. These extragrammatical factors give rise to alternants in which consonants more closely resemble their neighbours. The alternation itself is linguistically significant and thus represented in the grammar, but the resemblance is not.

## 6. SUMMARY

Both vowel harmony and consonantal lenition manipulate sound properties that have a linguistic marking function. In vowel harmony, the properties in question – the spectral correlates of vowel quality – are subject to grammatically represented assimilation. That is, their extension across suprasegmental spans is itself also linguistically significant, serving to demarcate morphosyntactic domains. In consonantal lenition, linguistically significant properties – such as noise bursts and abrupt amplitude changes – are

subject to grammar-internal suppression. Any assimilatory or coarticulatory accompaniments of this effect involve properties of the carrier signal and thus do not warrant representation in the phonological grammar.

## REFERENCES

- [1] P. Kiparsky, "Some consequences of Lexical Phonology", *Phonology* vol. 2, pp. 85-138, 1985.
- [2] F. Nolan, "The descriptive role of segments: evidence from assimilation", in *Papers in Laboratory Phonology II: Gesture, segment, prosody*, G. J. Docherty and D. R. Ladd, Eds., pp. 261-280. Cambridge: Cambridge University Press, 1992.
- [3] C. P. Browman and L. Goldstein, "Articulatory gestures as phonological units", *Phonology* vol. 6, pp. 201-252, 1989.
- [4] E. Flemming, "Scalar and categorical phenomena in a unified model of phonetics and phonology", *Phonology* vol. 18, pp. 7-44, 2001.
- [5] R. M. Kirchner, *An effort-based approach to consonant lenition*, PhD dissertation, UCLA, 1998.
- [6] W. Labov, M. Yaeger and R. Steiner, *A quantitative study of sound change in progress*, Philadelphia, PA: US Regional Survey, 1972.
- [7] P. Boersma, *Functional phonology*, The Hague: HAG, 1998.
- [8] H. Traunmüller, "Conventional, biological and environmental factors in speech communication: a modulation theory", *Phonetica* vol. 51, pp. 170-183, 1994.
- [9] J. J. Ohala and H. Kawasaki-Fukumori, "Alternatives to the sonority hierarchy for explaining segmental sequential constraints", in *Studies for Einar Haugen*, S. Eliasson and E. H. Jahr, Eds., pp. 343-365. Berlin: Mouton de Gruyter, 1967.
- [10] J. J. Ohala, "The phonetics and phonology of aspects of assimilation", in *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, J. Kingston and M. E. Beckman, Eds., pp. 258-275. Cambridge: Cambridge University Press, 1990.
- [11] L. A. Lavoie, *Phonological patterns and phonetic manifestations of consonant weakening*, PhD dissertation, Cornell University, 2000.
- [12] J. Harris and E-A. Urua, "Lenition degrades information: consonant allophony in Ibibio", *Speech, Hearing and Language: Work in Progress*, vol. 13, pp. 72-105. University College London, 2001.
- [13] P. Carr, *Phonology*, London: Macmillan, 1993.
- [14] G. Lindsey and J. Harris, "Phonetic interpretation in generative grammar", *UCL Working Papers in Linguistics*, vol. 2, pp. 355-369, 1990.