PHONOLOGY

ARTICLE

A unified model of lenition as modulation reduction: gauging consonant strength in Ibibio

John Harris,[1] Eno-Abasi Urua[2] and Kevin Tang [ID][3]

[1]Department of Linguistics, University College London, London, UK.
[2]Department of Linguistics and Nigerian Languages, University of Uyo, Uyo, Nigeria.
[3]Department of English Language and Linguistics, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany.

Corresponding author: John Harris; Email: john.harris@ucl.ac.uk

Abstract

We review and elaborate an account of consonantal strength that is founded on the model of speech as a modulated carrier signal. The stronger the consonant, the greater the modulation. Unlike approaches based on sonority or articulatory aperture, the account offers a uniform definition of the phonetic effect lenition has on consonants: All types of lenition (such as debuccalisation, spirantisation and vocalisation) reduce the extent to which a consonant modulates the carrier. To demonstrate the quantifiability of this account, we present an analysis of Ibibio, in which we investigate the effects of lenition on the amplitude, periodicity and temporal properties of consonants. We propose a method for integrating these different acoustic dimensions within an overall measure of modulation size. Not only does the modulated-carrier account cover all the classically recognised lenition types, but it also encompasses loss of plosive release in final stops – which, although not traditionally classed as lenition, is clearly related to processes that are.

Contents

1. Introduction

What impact do lenition processes have on the phonetic make-up of consonants? The very existence of a single term *lenition* suggests some kind of unitary effect. However, it is surprisingly hard to pin down what that effect might be, at least when we come at the issue from the viewpoint of what are perhaps the two best-known definitions of lenition: an increase in articulatory aperture or an increase in sonority.

   To see this, consider four classically recognised processes that fall under the rubric of lenition: spirantisation (e.g., *p > f*), vocalisation (*p > w*), debuccalisation (*p > ʔ*) and intervocalic obstruent voicing (*p > b*). Defining lenition as an increase in articulatory aperture draws together spirantisation, vocalisation and debuccalisation, but not intervocalic voicing. Defining lenition as an increase in sonority draws together spirantisation and vocalisation, but it is not clear how this applies to intervocalic voicing, or whether it applies at all to debuccalisation.

   In this article, we review and elaborate an alternative to these two models, one that we believe does provide a unitary phonetic account of lenition's impact on consonants. It is founded on the conception of speech as a modulated carrier signal (Dudley 1940; more references below). The carrier is a schwa-like signal that allows speech to be heard. The linguistic message is borne by acoustic events that modulate the carrier. From the perspective of this model, all types of lenition can be seen as reducing the extent to which a consonant modulates the carrier. The notion of lenition as modulation reduction has been explicitly developed by, among others, Harris & Urua (2001) and Harris (2003, 2004, 2009), and it is also implicit in earlier versions of element theory (Harris 1994; Harris & Lindsey 1995). A similar, albeit more specific notion is invoked in proposals that one type of lenition – spirantisation – reduces the extent to which a consonant interrupts the stream of speech (Kingston 2008) or causes auditory disruption in relation to neighbouring sounds (Katz 2016).

   We illustrate the modulated-carrier (MC) account with phonological and acoustic analyses of Ibibio, a Lower Cross language spoken in southeastern Nigeria. Lenition in Ibibio poses a number of descriptive challenges that are by no means unique to the language. We focus on two of them here. First, the outputs of lenition do not always fit easily into standard phonetic or feature classifications. Close auditory-impressionistic and instrumental inspection of spirantisation and vocalisation reveals a degree of continuous phonetic variability that potentially undermines the notion of a categorical distinction between lenited and unlenited consonants. This is an issue that is familiar from the analysis of spirantisation in well-studied languages such as Spanish. Second, spirantised/vocalised outputs show a close bond with unreleased stops in that they alternate with one another in prosodically weak positions. Spanish lacks anything like this phenomenon, though it is strongly reminiscent of the alternation between tapped and unreleased variants of /t/ in English (compare the tap in *get a* with the unreleased stop in *get by*). Loss or absence of release in stops is not traditionally classified as a lenition effect. However, the recurring link with spirantisation and vocalisation suggests that it should be.

   Both of these descriptive challenges can be successfully met by approaching them from the viewpoint of the MC model of speech. First, even in the face of continuous phonetic variation, the model allows us to draw a consistent distinction between lenited and unlenited consonants, even if not along the exact same lines as those drawn by traditional terminology. Second, the model unambiguously groups lack of plosive release alongside effects more traditionally recognised as lenition.

   Let us emphasise that this article is about the effect of lenition, not its cause.[1] Much of the research on lenition has concentrated on the latter question. According to a long-established assumption dating from the philological tradition, the sound changes that produce lenition are caused by articulatory undershoot in speech (see Hock 1991 and Honeybone 2008 for reviews of the relevant literature). There are important points of disagreement within this overall account. For example, is lenition-as-undershoot motivated by a pressure to minimise articulatory effort (e.g., Kirchner 1998, 2004) or not (e.g., Bauer 2008)? Is the effect of this pressure purely historical (e.g., Blevins 2004), or does it actively linger in synchronic grammars (e.g., Flemming 1995; Kirchner 1998)? Although the undershoot account posits

---

[1]Nor is this article about the factors that kick-start lenition in individual languages – the actuation issue addressed by Cohen Priva (2017).

a unitary articulatory cause of lenition, it is noticeable that the relevant literature generally steers clear of the question of whether there is also a unitary articulatory effect.[2] If undershoot is the root cause, we might expect any such effect to be covered by the definition of lenition as an increase in articulatory aperture. However, as we will see below, not all types of lenition fall within this definition.

Although our focus here is on the effect of lenition, this is not to say the MC model has nothing to contribute to the causation debate. It has been argued elsewhere that an MC-informed account of lenition offers a listener-oriented explanation of why grammars have it in the first place (Harris 2003, 2009). The account is part of a broader research development that focuses on the way lenition, in accentuating auditory-perceptual differences between strong and weak consonants, has the potential to provide listeners with cues for prosodic and morphosyntactic parsing (e.g., Harris 2004; Kingston 2008; Shiraishi 2009; Katz 2016).

The presentation runs as follows. We start in §2 with a review of the different types of consonant lenition and show how their segmental effects cannot be unified using approaches based on sonority or articulatory aperture. In §3, we outline the MC model of lenition. In §4, we provide a phonological description of Ibibio, focusing on the segmental details of lenition and the phonological environments within which it operates. In §5, we present an acoustic study of Ibibio lenition that is informed by the MC model. We conclude in §6 by considering the implications of this approach for the question of why grammars have lenition in the first place.

2. Consonant lenition

2.1. *The main lenition types*

To reaffirm that consonant lenition or weakening really is a unitary phenomenon, let us start by highlighting two properties that the different types have in common. One has to do with positional strength: In languages where lenition is sensitive to prosodic domain structure, it typically occurs in weak positions, such as non-initial in the word, foot or stem. The second has to do with segmental strength, which we can summarise by recalling Vennemann's often-cited definition, which Hyman (1975: 165) states as 'a segment X is [...] weaker than a segment Y if Y goes through an X stage on its way to zero'. The definition has proved extremely useful over the years, and we will invoke it again below. However, it says nothing about whether lenition might have a unitary phonetic effect on consonants. It allows for the possibility that progression towards zero can follow different phonetic trajectories with different phonetic halt points.

The main types of lenition can be broadly classified as spirantisation, vocalisation, obstruent voicing and debuccalisation.[3] Other terms used to describe lenition generally refer to processes that can be considered more specific instantiations of these main types. For example, tapping or flapping is a version of vocalisation, while glottalling is a version of debuccalisation. Most of these terms are applied interchangeably to both diachronic and synchronic manifestations of lenition, in acknowledgement of the fact that most synchronic lenitions are reflexes of identifiable sound changes. The main types are illustrated in (1) by brief examples from various dialects of Spanish. (For more extensive exemplification and references, see the cross-linguistic surveys in Kirchner 1998; Lavoie 2001; Gurevich 2004, 2011.)

---

[2] As Bauer (2008) acknowledges, defining lenition as articulatory undershoot draws in many types of assimilation. For example, place assimilation in *that man* [ðap man] reflects a failure to achieve an alveolar target. This extension is undesirable since assimilations do not fit the guiding definition of lenition as pushing a segment towards zero.

[3] Although degemination is often included in this list (e.g., Lass & Anderson 1975), we prefer to view it as a weight-based phenomenon rather than as a type of segmental lenition. The process can of course act as an entry point to lenition: Once shortened, a consonant is no longer protected by geminate inalterability (Hayes 1986) and is thus susceptible to lenition. We return to this point in our treatment of Ibibio below.

(1)   *Lenition in Spanish dialects*

    a.   *Gran Canaria: p, t, k > b, d, g > β, ð, ɣ*

        i.   tipi**k**o   tipi**g**o   'typical'

            una **t**jenda   una **d**jenda   'a shop'

            la **k**ama   la **g**ama   'the bed'

        ii. ro**b**a   ro**β**a   '(s)he steals'

            na**d**a   na**ð**a   'nothing'

            la **g**ana   la **ɣ**ana   'appetite'

    b.   *Cibao: l, r > j*

            sa**l**   sa**j**   'salt'

            ka**r**ta   ka**j**ta   'letter'

    c.   *Andalucia, Caribbean: s > h*

            a**s**ta   a**h**ta   'until'

            de**s**pwe**s**   de**h**pwe**h**   'after'

(1a) illustrates an intervocalic chain shift in Gran Canaria Spanish, where plain plosives undergo voicing (as in (1a-i)) while voiced plosives spirantise (as in (1a-ii)) (Trujillo 1980; Broś *et al*. 2021). The examples in (1b) illustrate vocalisation in Cibaeño Spanish, where coda liquids turn into glides (Harris 1969). (1c) shows debuccalisation of coda *s,* a characteristic of Spanish in Andalucia and the Caribbean (Harris 1969).

The spirantisation example in (1a) already touches on the descriptive challenge alluded to in §1. The names we give to the different types of lenition are largely inherited from the philological tradition. Although they have served well in establishing the broadest outlines of the phenomenon, they can be descriptively problematic when it comes to examining particular cases. The term *spirantisation* implies a prototypically spirant or fricative output. This is generally true where the output is voiceless (as in the part of the High German Consonant Shift that produced *p > f, t > s, k > x*). However, where spirantisation is accompanied by voicing, as in (1a), outputs broadly transcribed as *β, ð, ɣ* typically show variable degrees of articulatory stricture and are often more accurately described as frictionless continuants or approximants (cf. Kirchner 1998, ch. 4; Lavoie 2001). This is now well established for Spanish (Romero Gallego 1995; Cole *et al.* 1998; Ortega-Llebaria 2003; Kingston 2008; Figueroa Candia 2016; Broś *et al*. 2021), and we will see below that the same goes for Ibibio.

There is a plausible aerodynamic explanation for the preferred absence of frication noise in consonants that undergo both voicing and spirantisation: vocal fold vibration inhibits airflow, thus reducing the potential for air turbulence at the point of stricture (Ohala 1983). Terms such *approximantisation* (Broś *et al*. 2021) or *vocalisation* might be better here, if these are taken to mean the conversion of a consonant into a vocoid, a point we return to in our analysis of Ibibio. Nevertheless, the term spirantisation seems to have stuck, perhaps in recognition of the fact that vocoid outputs of lenition are not always canonical approximants like the yod output illustrated in (1b).

All of the processes illustrated in (1) meet Vennemann's definition of lenition. Cross-linguistic comparison and historical reconstruction allow us to identify the outputs of these processes as intermediate stages on trajectories that potentially culminate in consonant deletion (Lass & Anderson 1975). Note that the definition suggests we should also count loss of plosive release as a type of lenition: 'Applosives' are often observed to occupy an intermediate stage before stop debuccalisation, e.g., *t > tˀ > ʔ > ∅* (as in Middle Chinese to modern Mandarin; Chen 1976). We take up this point in our analysis of Ibibio below.

The fact that the different types of lenition share properties defined by positional and segmental strength admittedly does not entail that they must have something in common phonetically. Indeed, although the collection of main lenition types illustrated in (1) is relatively small, their phonetic effects

seem on the face of it pretty diverse. For example, spirantisation produces articulatory and aerodynamic conditions in the oral cavity that are quite different from those created by debuccalisation at the larynx. So it is perhaps unsurprising that neither of what are arguably the two best known accounts of lenition offers a unified phonetic characterisation of its effects. However, just because a phonetic unity has proved elusive is no reason to give up searching for one.

2.2. *Opening and sonority accounts of lenition*

Consider first an 'opening' account, in which lenition is defined as increasing the degree of articulatory aperture in a consonant, thereby reducing impedance to airflow through the vocal tract. This successfully brings together spirantisation and vocalisation (Lass & Anderson 1975; Broś *et al*. 2021), as well as debuccalisation, all of which result in a loosening of articulatory stricture. However, it excludes obstruent voicing. Compared with the open-glottis state of a voiceless consonant, the closed phase of vocal-fold vibration in a voiced congener increases impedance to airflow. Lass & Anderson (1975) are quite explicit that this discrepancy undermines any notion that lenition has a unitary articulatory impact. Instead, they propose two fundamentally separate types: opening proper and voicing (for the latter of which they prefer the traditional philological term *sonorisation*). The main problem here is that, in spite of their apparently opposite aerodynamic effects, voicing and spirantisation are closely related. They typically occur in the same phonological contexts, and one often accompanies or precedes the other (as in the chain shift illustrated in (1a)).

Now consider the definition of lenition as increasing the sonority of a consonant (e.g., Lavoie 2001; Smith 2008). This proposal is likely to be a non-starter for any researcher who questions the phonetic credentials of sonority (e.g., Ohala & Kawasaki-Fukumori 1997). However, let us for the sake of argument assume what is perhaps the most widely claimed (if not the only) ingredient in phonetic definitions of sonority, intensity (e.g., Hankamer & Aissen 1974; Parker 2002; Albert & Nicenboim 2022). Spirantisation and vocalisation certainly have the measurable effect of increasing the intensity of a consonant. This has been well demonstrated for Spanish (see the references above), and again we will see that it also holds true of Ibibio. However, an increase in intensity is rather more difficult to show for debuccalisation and stop voicing.

To see this, consider the impact lenition has on the intensity properties of oral stops. One effect of stop debuccalisation is the loss of the high-intensity noise burst associated with plosive release. On the face of it, this reduces the intensity of a consonant rather than increases it. A similar problem crops up with lenition by voicing, and in particular with whether we gauge the sonority of an oral stop on the basis of its closure phase, its release phase or the two combined. The low-frequency energy associated with the closure phase of a voiced stop makes this part of it more intense than a voiceless counterpart. However, the reduced airflow across the glottis that accompanies vocal-fold vibration means that the release burst of a voiced stop is actually less intense than that of a voiceless congener.

Both of these issues might be addressed by incorporating the dimension of time into the definition of sonority: The intensity of a more sonorous segment is not just higher but also prolonged (Parker 2002; Clements 2009; Albert & Nicenboim 2022). Depending on how this definition is formulated, the release burst of a plosive may or may not be considered too brief to count towards its sonority value. One way of including it is to measure the intensity of a stop as an average of the minimum in the hold phase and the maximum in the burst (Parker 2011). However, even with this revision, oral stops and glottal stop end up occupying the same niche on the sonority hierarchy (see the detailed hierarchy in Parker 2011). That being the case, debuccalisation still cannot be characterised as an increase in sonority.

Lavoie (2001) herself acknowledges this point by characterising voicing and stop debuccalisation as decreasing markedness rather than increasing sonority. However, the formal device of markedness is simply a way of describing cross-linguistic phonological preferences and does not embody a phonetic dimension along which lenition operates, let alone one that can be subsumed under sonority.

A further problem for the sonority account of lenition concerns the niche occupied by nasals on the sonority hierarchy. Consonants entering a trajectory involving more than one type of lenition (such as

the $p > b > \beta$ chain illustrated in (1a)) are observed to proceed one step at a time (Lass & Anderson 1975; Foley 1977). The location of nasals between fricatives and approximants on the sonority scale wrongly predicts that vocalising obstruents should pass through a nasal stage on their way to becoming approximants (e.g., *$p > m > w$). What we find instead is that vocalising oral consonants remain consistently oral (Szigetvári 2008).

This is not to downplay the phonetic impact lenition can have on the intensity of consonants. Indeed, it is one of the acoustic dimensions we investigate in our analysis of Ibibio below. However, two points should be clear. First, rather than increasing intensity as per the sonority account, certain types of lenition either decrease it or at least fail to alter it. Second, intensity on its own does not offer a full picture of lenition's phonetic impact. Sonority was originally proposed as a way of describing syllabification and phonotactics. Stretching it to a use for which it was not designed, in an attempt to unify various types of lenition, proves to be a step too far.

This latter point chimes with one of the criticisms levelled at sonority accounts of consonant-cluster phonotactics: Constraints based on sonority or intensity sequencing are not sufficient to explain all of the recurrent patterns we observe in clustering behaviour, such as the place restrictions that disfavour complex onsets such as *pw, bw, tl, dl.* Phonotactic patterns are better explained by differences in the recoverability of consonants' auditory–acoustic cues in different phonological positions (Wright 2001, 2004). While intensity is undoubtedly one of the factors involved here (the louder the cue, the more audible it is), it is by no means the only one. A crucial role is also played by the degree to which the speech signal is modulated across adjacent segments: The greater the modulation, the more readily discriminable are the consonants (Ohala & Kawasaki-Fukumori 1997; Harris 2006; Henke *et al.* 2012). The notion of modulation size will also figure in the account of lenition to be presented below, with the difference that size is to be gauged between a consonant and the carrier signal rather than between one consonant and the next.

To summarise: accounts based on sonority or articulatory opening fail to provide a unitary phonetic definition of lenition. In light of this, it is understandable that some researchers have concluded that lenition is not a single phenomenon after all (Lass & Anderson 1975; Lavoie 2001; Katz 2016). However, we will now turn to an alternative account under which lenition can indeed be seen to have a unitary phonetic impact on consonants.

3. Lenition as modulation reduction

3.1. Speech as a modulated carrier signal

The model of lenition we review here is based on the characterisation of speech as a modulated carrier signal (Dudley 1940; Traunmüller 1994, 2005; Ohala & Kawasaki-Fukumori 1997). The carrier is linguistically void, furnishing information that reveals details about the talker's organism (sex, age, size, etc.), attitude or emotional state, and physical location. Modulations of the carrier contain the linguistic information that conveys lexical–grammatical meaning. In short, the carrier allows the linguistic message to be heard, while modulations allow the message to be understood.

An unmodulated carrier signal is the schwa-like acoustic effect produced by a neutrally open vocal tract. It lacks the spectral prominences that result from the convergence of formants, and it is typically periodic (though not necessarily so – see, e.g., Traunmüller 2005 on whispering). Modulations deviate from this baseline along the parameters of amplitude, spectral shape, periodicity, fundamental frequency and rate of change. The magnitude of a modulation can be defined as the length of the trajectory it traces through a multidimensional acoustic space defined by these parameters (Ohala & Kawasaki-Fukumori 1997).[4]

---

[4]Coleman *et al.* (2015) employ a similar notion of multidimensional acoustic space to model sound change.
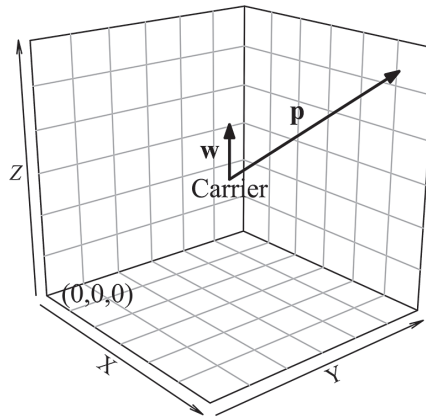
Figure 1. *Schematic illustration of the difference in modulation size between* p *and* w *(after Ohala & Kawasaki-Fukumori* 1997*). The axes (X, Y, and Z) represent three different acoustic parameters. The carrier signal lies at the midpoint of the cube. The trajectory followed by* p *away from the carrier is longer than that followed by* w.

3.2. The modulated-carrier model of lenition

To gain an initial idea of how the MC model applies to lenition, consider the ways different consonants modulate the carrier signal when in intervocalic position. For the purposes of the illustration given in (2), the flanking vowels are schwa – the quality of the bare carrier – and the dimensions of amplitude and rate are collapsed. The consonants in (2b–f) feature here because they make up a pretty much exhaustive collection of possible outcomes of lenition of *p*. However, the comparisons to be drawn among them below hold regardless of whether the consonants result from lenition: For example, [w] modulates the carrier in a particular way regardless of whether it is a reflex of /w/ or the result of /p/ > [w].

(2) Modulation	a. əpə	b. əwə	c. əbə	d. əfə	e. əʔə	f. əhə
Change in spectral shape	✓	✓	✓	✓	✗	✗
Radical drop in amplitude	✓	✗	✓	✗	✓	✗
Aperiodicity	✓	✗	✓	✓	✗	✓
F0 discontinuity	✓	✗	✗	✓	✓	✓

Compare first the signal characteristics of a voiceless labial plosive in (2a) with those of a labial approximant in (2b), such as we witness in *p* > *w* vocalisation. In the case of the plosive, the carrier is modulated along each of the acoustic parameters just mentioned. There are abrupt radical changes in amplitude at the onset and offset of the plosive. There is a burst of aperiodic energy at the point of release. There are rapid spectral transitions at the onset and offset. Periodicity – and with it fundamental frequency – is discontinued during the plosive. In the case of the approximant, modulation of the carrier consists almost entirely of a change in spectral shape, taking the form of smooth formant transitions between the glide and the surrounding vowels. Any amplitude change that might occur during the glide is much less marked and abrupt than in the case of the plosive. The modulation produced by the plosive is thus larger than that produced by the glide: In engaging all of the acoustic parameters in (2), the plosive travels a longer distance through multidimensional acoustic space than the glide. The difference is schematised in Figure 1.

The notion of modulation size provides us with a straightforward definition of consonantal strength: The stronger the consonant, the larger the modulation. This in turn allows us to define lenition as

having a unitary phonetic impact on its targets: It reduces the extent to which a consonant modulates the carrier signal.

This definition holds not just for vocalisation (as in (2b)) but also for obstruent voicing (2c), spirantisation (2d), stop debuccalisation (2e) and spirant debuccalisation in (2f). What distinguishes one type of lenition process from another is the particular combination of acoustic parameters along which this overall reduction occurs. In the case of spirantisation, vocalisation and voicing, the relevant parameters are amplitude and periodicity. In the case of debuccalisation, they are primarily periodicity and spectral shape. Put differently, to varying extents, all of these processes push a consonant in the direction of merger with the carrier signal. If historical lenition is allowed to culminate in the deletion of a consonant, the result is total merger with the carrier. This is how the MC model enacts Vennemann's definition of lenition as progression towards zero.

The notion of merger with the carrier is similar in spirit to the observation that lenition reduces the extent to which a consonant perturbs intensity across a VCV sequence (Kingston 2008; Kaplan 2011; Katz 2016). As we have seen, defining lenition as increasing intensity only applies to spirantisation and vocalisation. Katz (2016) suggests a difference in kind between this type of lenition (what he calls 'continuity lenition') and debuccalisation ('loss lenition'). In contrast, the MC model unequivocally unites all of these processes, because it encompasses not just intensity but also the other acoustic parameters in (2). Pushing a consonant closer to the carrier can be achieved not just by suppressing an abrupt drop in intensity (as in vocalisation) but also by, for example, suppressing aperiodic energy and a change in spectral shape (as in stop debuccalisation). Moreover, as we will see below, the MC model allows us to extend this notion of merger with the carrier to environments other than intervocalic position.

When intervocalic lenition produces voiced outputs, it takes the form of obstruent voicing (2c) or vocalisation (2b), often in historical sequence (e.g., *p > b > w*). With intervocalic stops, the effect is attributable to the spontaneous voicing of the surrounding vowels being interpolated passively through the intervening consonant (Westbury & Keating 1986; Rice & Avery 1989; Kirchner 1998; Avery & Idsardi 2001; Iverson & Salmons 2003; Harris 2004; Jansen 2004). In the case of vocalisation, the loosening of the oral stricture creates more favourable aerodynamic conditions for vocal-fold vibration. From an MC perspective, both of these processes reveal the carrier seeping throughout the entire VCV sequence (cf. Harris 2009; Botma & van 't Veer 2013).

It might be objected that, in pressing the case for a unified definition of lenition, we lose sight of what Katz (2016) claims to be an important distinction between his two types of lenition: The loss type 'frequently' triggers positional neutralisation, while the continuity type 'very rarely' does so. However, it is not clear how strong this correlation is or even whether it exists at all (see the survey in Gurevich 2004). It is certainly true that intervocalic spirantisation/vocalisation in Spanish, perhaps the best-known example of continuity lenition, does not result in neutralisation. However, the same phenomenon in Ibibio does: Here, as we will see below, vocalisation neutralises a laryngeal contrast that holds between stops in stem-initial position. English provides evidence that runs doubly counter to the supposed correlation. The continuity process of intervocalic tapping neutralises the contrast between /t/ and /d/ (producing homophony between *latter* and *ladder*). On the other hand, the loss process of glottalling in British English maintains the /t/–/d/ contrast: Glottal stop is the unique allophone of /t/ (hence [ʔ] in *latter, bat* but [d] in *ladder, bad*). In any event, a unified definition of lenition does not prevent us from reporting whether specific cases trigger neutralisation or not.

Our focus in this article is on the set of processes that are generally agreed to fall under the rubric of lenition. However, the MC model of segmental strength extends to processes for which a connection with lenition is not so generally agreed upon. For example, it unambiguously classifies final obstruent devoicing – alongside suppression of stop release – as weakening (Harris 2009), *pace* its traditional description as hardening. Another example is the apparently diverse collection of processes that constitute vowel reduction. There are good grounds for treating all of these as forms of segmental weakening (e.g., Barnes 2006). All of them can be understood as reducing in the extent to which a vowel modulates the carrier (Harris 2005).

4. Lenition in Ibibio

In this section, we outline the main facts pertaining to lenition in Ibibio. To understand the environments under which it operates, we need to start with the language's verb morphology, which is built around an inflectional stem consisting of a root and an optional suffix. An areal characteristic of the genetically diverse languages in the Nigeria–Cameroon border region is that the shape and size of the stem are subject to quite severe prosodic-templatic restrictions (see Hyman 1990; Hyman *et al.* 2019).[5] In Ibibio, the basic template takes the form $C_1V_1XC_2V_2$, where X is a copy of either V_1 (CVVCV) or C_2 (CVCCV). The resemblance to a heavy-light trochee is unmistakable, as has been pointed out by Urua (1990, 2000) and Akinlabi & Urua (1992, 1994). Similar observations have been made about the closely related language Efik (Cook 1985; Hyman 1990).

As a lexical-tone language, Ibibio lacks stress prominence, the property most usually associated with the metrical foot. Nevertheless, the CVXCV template does display weight and segmental asymmetries that are strongly reminiscent of trochees in stress languages. For example, the second syllable of the template must be light, while the first can and, in certain paradigms, must be heavy. A full range of vowel contrasts is supported in the first syllable but not in the second. Lenition, as we will see presently, occurs in stem positions that correspond to weak positions in metrical trochees.

It is debatable whether these facts indicate that Ibibio actually has metrical feet. The stem template might instead reflect a universally preferred prosodic shape for root+suffix combinations (Downing 2006). This issue is not immediately relevant to our present concerns, however, and below we will simply refer to the stem.

The CVXCV template in Ibibio places an upper bound on the size of the stem. (3) illustrates the six attested canonical stem shapes that are contained within this limit.

(3) a. CV   dí   sé   da
                 'come'   'look'   'stand'

     b. CVC   kàt   tɔ́k   tóp
                 'show'   'urinate'   'throw'

     c. CVCV   kárá   séɣé   bɔ́ɣɔ́
                 'govern' 'be childish' 'overtake'

     d. CVVC   dííp   kóót   bùùk
                 'hide'   'call'   'bury'

     e. CVCCV   bèkké   bɔ̀kkɔ́   bàkkɔ́
                 'belch'   'escape'   'uproot'

     f. CVVCV   sèèmé   tòòró   sɔ̀ɔ̀ŋɔ́
                 'bemoan' 'praise'   'affirm'

The template also acts as a lower bound for certain verbal paradigms. In this case, potentially oversized or undersized morphological material is tailored to a fixed CVVCV or CVCCV template through segment truncation or augmentation.[6] These effects are illustrated by the negative, frequentative and reversive forms in (4)–(6). (In the examples given here and below, stems are marked out by square brackets.)

(4) *Negative: root* + ké

     a. *CVC*

         [tèm] 'cook'   í[tèmmé] 'is not cooking'

---

[5] For detailed treatments of foot-templatic effects in other languages in the same geographical area as Ibibio, see, for example, Hyman (1982, 2003, 2008) on Gokana, Basaa and Northwest Bantu (respectively).

[6] As with spirantisation and vocalisation discussed above, the use of terms such as truncation and augmentation is not intended to imply a derivational analysis. In keeping with a fully output-oriented approach, they can be understood here in a purely descriptive sense to refer to cross-paradigm comparisons.

[kɔ̀p]	'lock'	í[kɔ̀ppɔ́]	'is not locking'
[fʌ́k]	'cover'	í[fʌ́kkɔ́]	'is not covering'

    b. *CV*

[sé]	'look'	í[sééɣé]	'is not looking'
[nɔ̀]	'give'	í[nɔ̀ɔ̀ɣɔ́]	'is not giving'
[dá]	'stand'	í[dááɣá]	'is not standing'

    c. *CVVC*

[síít]	'block'	í[sííɾé]	'is not blocking'
[fáák]	'wedge'	í[fááɣá]	'is not wedging'
[kɔ́ɔ́ŋ]	'hang on hook'	í[kɔ́ɔ́ŋɔ́]	'is not hooking'

(5) *Frequentative: root + N*é

    CV

[nɔ̀]	'give'	[nɔ̀ŋŋɔ̀]	'give (freq.)'
[k͡pá]	'die'	[k͡páŋŋá]	'die (freq.)'

(6) *Rerversive: root + C*é

    CVVC

[síít]	'block'	[sítté]	'unblock'
[fáák]	'wedge'	[fákká]	'remove wedge'
[kɔ́ɔ́ŋ]	'hang on hook'	[kɔ́ŋŋɔ́]	'unhook'

Stems consisting of a CVC root and a CV suffix satisfy the fixed-templatic restriction by default (see (4a)). Attachment of a CV suffix to CV roots is accompanied by either vowel lengthening (see (4b)) or consonant gemination (see (5)). Suffixation to CVVC roots results either in consonant truncation (see (4c)) or vowel shortening (see (6)).

The segmental asymmetries evident within the Ibibio stem are a specific instantiation of the more general pattern of relative positional strength or prominence (Beckman 1999; Hyman 2008). For our immediate purposes, it will be useful to draw a terminological distinction between a strong *head* sector, consisting of the initial CV of the template, and a weak *tail*, consisting of any residual positions. Only the head sponsors the full set of vowel and consonant distinctions in Ibibio. The contrastive potential of the tail is greatly curtailed: Not only does it lack a proportion of the segmental material available to the head, but what material it does have is to a great extent assimilated from the head.

In the case of vowels, assimilatory neutralisation in the tail manifests itself in two ways. First, as indicated by the examples in (4b), (4c) and (6), V_2 in a V_1V_2 cluster is invariably a copy of V_1, both tonally and segmentally. Second, a stem-final vowel is harmonically dependent on the head nucleus. This effect, already suggested by some of the examples above, is more fully exemplified by the frequentative and relative forms in (7). Here we see how the head nucleus determines height, frontness and (for back vowels) ATR in a stem-final vowel.

(7) a.

[díí-mé]	'lift up'	[kúú-mɔ́]	'open'
[fèè-ŋé]	'run'	[dóó-mó]	'become heavy'
[k͡páŋ-ŋá]	'die'	[dɔ́ɔ́-mɔ́]	'wipe up'

    b.

áà[sììrè]	'who blocks'	áà[dùùɣò]	'who is alive'
áà[sèèɣè]	'who looks'	áà[dòmmò]	'who bites'
áà[dààɣà]	'who stands'	áà[kɔ́ppɔ́]	'who locks'

Consonant positions within the stem also exhibit asymmetries, and it is here that we witness the effects of lenition. The contrastive asymmetries between head and tail consonants are broadly summarised in (8), which shows the distribution of oral stops and related segments.

(8)

STEM HEAD		STEM TAIL		
[C		VC:V	VC]C	VC(])V
k͡p	b	pp	p	β
t	d	tt	t	ɾ
k		kk	k	ɣ

Consonants in the tail are subject to neutralisation, failing to sustain a proportion of the laryngeal and manner contrasts borne by the syllable onset of the head. (Here we see the neutralising effect of 'continuity' vocalisation in Ibibio, mentioned in §3.2.) The laryngeal contrast holding in head onsets takes the form of a distinction between plain (voiceless unaspirated) and prevoiced plosives (see Urua 1990; Connell 1991; Harris 2004). In tail-internal geminates, this contrast is suspended in favour of plain stops.

It is singleton consonants within the tail that are targeted by lenition. Geminates are immune, exhibiting characteristic inalterability (cf. Hayes 1986). The effect of lenition on singletons is illustrated in (9), where the transcriptions suggest what would traditionally be called spirantisation (*p > β, k > ɣ*) and vocalisation (*t > ɾ*). For the time being, these broad terms and symbols will serve as placeholders until we examine the phonetics of lenition in more detail below.

(9)   a.   [díp]   'hide'   [díβé]   'hide oneself'
         [déép]   'scratch'   [dééβé]   'not scratching'
         [bɔ́p]   'tie'   [bɔ́βɔ́]   'tie oneself'

         [bèt]   'shut'   [bèɾé]   'be shut'
         [kóót]   'call'   [kóóɾó]   'not calling'
         [síít]   'block an opening'   [síɾé]   'be blocked'

         [fʌ́k]   'cover'   [fʌ́ɣɔ́]   'cover oneself'
         [fáák]   'wedge'   [fááɣá]   'not wedged'
         [bók]   'group together'   [bóɣó]   'be grouped together'

   b.   [kɔ̀p]   'lock'   [kɔ̀β] úsʌ́ŋ   'lock the door'
         [bèt]   'push'   [bèɾ] ówó   'push someone'
         [kʌ̀k]   'shut'   [kʌ̀ɣ] úsʌ́ŋ   'shut the door'

The alternating pairs illustrated in (9) show the only singleton oral consonants that are permitted within the tail in Ibibio. (Nasal stops, which are immune to lenition, are also allowed in this position, e.g., *kpàn* 'debar', *sàŋá* 'walk'.) Utterance-finally or before a word-initial consonant, the segments are realised as stops. Impressionistically described, they tend to be unreleased and characterised by rapid decrescendo voicing from the preceding vowel (Urua 1990; Connell 1991). Intervocalically, they spirantise or vocalise. As a comparison of (9a) and (9b) shows, the triggering vowel may or may not be separated from the target consonant by a word boundary.[7]

---

[7]In the absence of an intervocalic contrast between geminate and non-geminate plosives here, it might initially seem reasonable to treat the plosives as singleton consonants (cf. Connell 1991). However, this would disturb the otherwise uniform quantitative

To lenite, an Ibibio singleton consonant must appear in the tail.[8] The necessity of this condition is confirmed by the fact that the presence of a following vowel is not in itself sufficient for lenition to take place. An intervocalic consonant resists lenition if it either occupies a head position or falls outside the stem. In the examples in (10), the prevocalic context is provided by a prefix vowel, nominalising in (10a), pronominal in (10b). The root-initial consonants following the prefix are in a head position and thus resist lenition.

(10) a. ú[táŋ]   *úɾáŋ   'plaiting'
          ú[kʌ́p]   *úɣʌ́p   'covering'

      b. í[tòòró]   *íɾòòró   '(s)he is praising'
          í[kàttá]   *íɣàttá   '(s)he is not showing'

To see what happens to consonants that are intervocalic but lie outside a stem, consider the examples containing the negative/reversive suffix -*ké* in (11). In (11a), the suffix is enclosed within a stem; the suffix consonant thus occupies a tail and undergoes lenition in the expected manner.

(11) a. [séé-ɣé]   'not look'
          [dáá-ɣá]   'not stand'

      b. [dáápá]ké   *dáápáɣá   'not dream'
          [kóŋŋɔ́]ké   *kóŋŋɔ́ɣá   'not unhook'
          [dámmá]ké   *dámmáɣá   'is not crazy'
          [fáŋŋá]ké   *fáŋŋáɣá   'not argue'

In the examples in (11b), on the other hand, the same suffix appears outside a stem that is already saturated by CVXCV material. Excluded from the stem, the suffix consonant is not subject to lenition. Note that vowel harmony too is stem-sensitive: As the examples in (11b) demonstrate, a stem-external suffix vowel does not harmonise with a root vowel.

The strongly foot-like behaviour of the Ibibio stem is underlined by the observation that the domain sensitivity of lenition illustrated in (11) is, aside from stress prominence, in all significant respects identical to what can be found in languages with stress feet. Focusing on coronals, compare the conditions on prevocalic tapping in Ibibio with those in English (square brackets here mark feet in English and stems in Ibibio):

(12)   English   Ibibio
      a. Stop: re[*táin*]   ú[táŋ]   'plaiting'
      b. Tap: [*cíty*]   [bèɾé]   'be shut'
      c. Tap: [*gèt*] *óff*   [βèɾ] ówó   'push someone'
              [*gét a*]

In English, stress is only tangentially implicated in tapping (despite what is often assumed; cf. Ladefoged 2001: 58). As the English examples in (12c) demonstrate, word-final /t/ taps regardless of whether a following vowel bears stress or not. What is crucial is the consonant's location in relation to foot structure: /t/ resists tapping when initial in the foot (as in (12a)) but undergoes it when non-initial (as in (12b) and (12c)) (see Kiparsky 1979; Harris 1998, 2013; Jensen 2000). The Ibibio stem exhibits precisely the same pattern.

---

patterning of verbal paradigms where we have independent evidence of a fixed heavy–light template. In the case of the negative paradigm, for example, a CVCV analysis of forms containing an intervocalic plosive would sever the quantitative link with clear cases of CVXCV involving either a long vowel (as in *kóóró* 'not call') or a geminate nasal (as in *dómmó* 'not bite'). Geminate nasals contrast with singletons intervocalically, e.g., *dómmó* 'not bite' *vs. dómó* 'switch on light'.

    [8]On reasons for rejecting the view that intervocalic leniting consonants are ambisyllabic in Ibibio, or indeed in any language, see Harris (2013).

When it comes to detailing the phonetics of lenition in Ibibio, we immediately face the general descriptive challenge mentioned at the outset of this article: It is not always easy to characterise the lenited reflexes in terms of traditional impressionistic articulatory labels. The broad transcriptions used in the Ibibio examples above gloss over continuous phonetic variability in the degree of stricture and frication noise observed in the lenited consonants. In Ibibio, this variability is to be found both dialectally and within the speech of individual speakers. Connell's meticulous phonetic description of the language speaks eloquently of the difficulties inherent in trying to apply standard terminology here: He refers variously to 'tapped approximants', 'tapped fricatives', 'tapped stops', 'approximant-like quality' and 'weak, unstable articulation' (Connell 1991: 65 ff.).

Impressionistically, it is not clear that individual variants transcribable as *β* or *γ* in Ibibio qualify as full-blown fricatives. As in Spanish, they do not exhibit the consistent level of frication noise we associate with prototypical fricatives. It is probably true to say that all languages possessing genuine voiced fricatives (French and Polish, for example) also have homorganic voiceless counterparts (cf. Ladefoged & Maddieson 1996, ch. 5). In contrast, where so-called spirantisation gives rise to a voiced continuant series, this is not necessarily matched by an existing homorganic voiceless series. This is the case in Ibibio, where there is no voiceless series corresponding to *β–r–γ*.

One of the aims of the instrumental study presented in the next section is to provide a more detailed description of the continuous effects of Ibibio lenition than is possible with impressionistic phonetic labelling.

5. Ibibio lenition and the speech signal

5.1. *Measuring modulation size*

In this section, we attempt a proof-of-concept demonstration of the MC model of lenition. The demonstration takes the form of an instrumental study of Ibibio that examines the acoustic respects in which consonants in lenition-favouring positions differ from those in lenition-resistant positions. The study seeks to answer the following questions: (a) Can the notion of modulation size be measured and, if so, (b) does the measure reliably distinguish different positions within the Ibibio stem? The study can be thought of as a first step towards identifying the auditory–acoustic cues Ibibio listeners rely on to make these distinctions.

The conventional classifications we employed in describing Ibibio lenition in §4 lead us to expect two acoustic dimensions to be most relevant to the task of establishing whether modulation size correlates with stem position in the language: intensity and periodicity. We first present the results of separate investigations into how the data generated by the study can be classified on the basis of these two dimensions. We label these investigations 'Edge' and 'Noise' (echoing the names of phonological features proposed by Harris & Lindsey 1995). We then go on to investigate how these separate measures can be integrated to provide an overall measure of modulation size.

The data are drawn from audio recordings of four adult native speakers of Ibibio (two female).[9] The subjects produced a set of words containing all the stops and related reflexes shown in (8), located within a range of phonological contexts representing different stem positions. The data consist of 504 spoken word tokens, equally distributed across the four speakers. For each word token, amplitude and periodicity measurements were taken within a VCX analysis frame, where C is a target consonant (singleton or geminate) preceded by a vowel and followed by a sonorant (vowel or nasal consonant). In the case of a word-final target consonant, the sonorant was supplied by a following word in a carrier phrase.[10] The carrier signal is of course not independently represented in conventional spectrographic

---

[9]The speech data on which this section is based were originally collected for a study reported by Harris & Urua (2001). Acoustic analysis of these data was performed using the SFS software designed by Mark Huckvale at University College London (https://www.phon.ucl.ac.uk/resource/sfs/).

[10]The carrier phrase is *ḿbô* ___ *ǹnɔ̀* 'I say ___ for myself'.

analysis. Nevertheless, having flanking sonorants within our analysis frame allows us to approximate the carrier signal as closely as is feasible in natural speech. The frame allows us to gauge the extent to which a target consonant deviates from surrounding segments that approximate the carrier.

Measurement within the analysis frame commenced at the midpoint of the pre-target vowel and ended at the midpoint of the post-target sonorant. As shown in (13), five stem positions, differentiated on the basis of the morphological paradigm criteria discussed in §4, are represented in the data.

(13)   Stem position   Example
    a. Initial   V[**C**V   ì-**t**à   'punch'
    b. Medial geminate   V**C**:V   sá**kk**á   'be split'
    c. Final preconsonantal   V**C**]C   tó**p** ǹnɔ̀   '…throw for me'
    d. Medial singleton   V**C**V   tá**β**á   'forfeit'
    e. Final prevocalic   V**C**]V   sá**r** á   'discriminate'

5.2. Amplitude: 'Edge'

5.2.1. Edge: method

As described in §3, one of the hypothesised effects of lenition is a reduction in the extent to which a consonant modulates the carrier along the dimension of intensity. Various methods have been proposed elsewhere for measuring the effect of lenition on the intensity properties of consonants, especially in Spanish. One method measures the difference in intensity between a consonant and a neighbouring vowel (Soler & Romero 1999; Lavoie 2001; Ortega-Llebaria 2003; Colantoni & Marinescu 2010; Hualde *et al.* 2010, 2011; Carrasco *et al.* 2012; Broś *et al.* 2021). Another measures the velocity with which intensity increases between the consonant and a following vowel, motivated by the observation that C-to-V transitions are more abrupt when the consonant is a stop than when it is an approximant (Kingston 2008; Hualde *et al.* 2010, 2011; Ennever *et al.* 2017). (For a useful comparison of these different methods, applied to Chilean Spanish, see Figueroa Candia 2016.)

Common to all of these methods is an analysis window that spans not just the target consonant but also the transitions in one or both of the flanking vowels. A very different approach is taken in a recent study by Tang *et al.* (2023), which quantifies the extent to which consonants are subject to lenition in a corpus of spoken Argentinian Spanish. The acoustic analysis window is limited to the target consonant itself, which is automatically classified in terms of binary values of [±sonorant] and [±continuant], two of the main features affected by spirantisation. The classification is based on a probabilistic comparison of the consonant's acoustic properties with those of reference segments described elsewhere in the literature.

The method we employ here is related to the transition-based studies above. However, it differs significantly in several respects, and not just because it is motivated by our specific aim of characterising lenition as modulation reduction. It also takes account of the phonological differences between lenition in Ibibio and Spanish. For one thing, we need an intensity measure that does not depend on the presence of a following vowel. Recall that, unlike Spanish, Ibibio has a series of word-final voiceless stops that are potential lenition targets. These consonants behave differently with respect to lenition according to whether they are followed by a vowel or a consonant in the next word (see again the examples in (9)). For another thing, also unlike in Spanish, we need a method that allows us to investigate the effect of stem position on lenition.

The basic method, which we call 'Edge', measures the fluctuation of acoustic energy across the VCX analysis frame described above. The method samples amplitude, in dB, at 10 ms intervals within the VCX frame and computes the standard deviation of these measures. (Below we take up the issue of whether this value should then be normalised by the duration of the frame.) The higher the Edge value, the greater the degree to which amplitude varies across the frame. Since the segments on both sides of the target consonant are high-energy sonorants, a relatively high Edge value can reasonably be taken to

reflect a drop in energy during the target consonant. (No tokens showed evidence of an energy increase during the target segment.)

By its very nature, a standard deviation metric takes no account of the temporal order of the values it samples. This places the metric at an advantage over methods that depend on the consecutive order of values (t minus $t − 1$), such as the C-to-V velocity measure employed by Kingston (2008) and Hualde *et al.* (2010, 2011) for Spanish. As Kingston (2008: 20) notes of his study, around 9% of data tokens had to be discarded because the extracted minimum and maximum intensity-change values were temporally misordered. No such data loss is necessary with a standard deviation metric.

The changes in energy we are interested in capturing constitute acoustic landmarks – areas of the speech signal where the correlates of consonant contrasts are most salient (Stevens 1985). Since the relevant landmarks are found at different frequencies, the Edge measure needs to be replicated across different frequency bands. We have a rough initial idea of what the relevant bands should be. For example, transitions for plosives (i.e., unlenited consonants) are best detected in the range of 2–5 kHz, while those for sonorants (the outputs of vocalisation) occur mainly in the range of 0.8–5 kHz (Liu 1996). We know rather less about whether the more specific bands used for well-studied languages such as English are valid for all languages, including less-studied ones such as Ibibio. In their study of Gurindji, Ennever *et al.* (2017) investigated the sensitivity of their measures of lenition (duration, magnitude of change of intensity and peak intensity velocity) to different frequency-band settings. They found that, particularly in the range of 400–1,100 Hz (±100 Hz), small changes to the precise band settings had little impact on their results.

Here we will compare two different band divisions: the one used by Liu (1996) for English and by Kingston (2008) for Spanish (LK) and the one used by Harris & Urua (2001) for Ibibio (HU).[11]

(14) a. Liu (1996) and Kingston (2008) (LK): 0–400, 800–1,500, 1,200–2,000, 2,000–3,500, 3,500–5,000, 5,000–8,000 Hz

   b. Harris & Urua (2001) (HU): 100–2,000 (low), 1,500–3,500 (mid), 3,000–5,000 (high), 100–5,000 (overall) Hz

5.2.2. Edge: results

Let us start by investigating the extent to which the two frequency bandings shown in (14) are correlated. Figure 2 depicts a hierarchical cluster analysis that allows us to compare how the two bandings classify the Edge data.[12] Three major clusters of frequency bands can be seen at a Pearson similarity cut-off threshold of 0.6. A lower or overall frequency band covers LK's 0–400 Hz and HU's 100–2,000 / 100–5,000 Hz. A mid band covers LK's 800–1,500 / 1,200–2,000 / 2,000–3,500 Hz and HU's 1,500–3,500 Hz. A high-frequency band covers LK's 3,500–5,000 / 5,000–8,000 Hz and HU's 3,000–5,000 Hz. This clustering suggests that the two bandings are highly correlated and thus that choosing between them is unlikely to make much of a quantitative difference – echoing the results of Ennever *et al.* (2017). In what follows, we will continue with HU in (14b), the banding adopted by Harris & Urua (2001).

The results in Figure 3 show the distribution of Edge values across the five stem positions, for all speakers and all frequency bands. (A breakdown of these values by speaker can be found in Figure A1 in Appendix A.) Based on the impressionistic description summarised in (8), we expect a two-way partition of the Edge data by stem position: lenition-resistant V[CV, VC:V, VC]C *vs.* lenition-prone VCV, VC]V. This expectation is broadly borne out by the distributions in Figure 3. Looking at the mean values (red dots), we can see that the stem positions fall into two broad clusters: The values for VCV, VC]V are consistently lower than those for V[CV, VC:V, VC]C. The lower values confirm that consonants in positions that are susceptible to vocalisation/spirantisation perturb amplitude less than consonants in positions that resist it.

---

[11]In the banding used by Harris & Urua (2001), no amplitude measurements were taken below 100 Hz, in order to exclude low-frequency rumble.

[12]The cluster analysis was performed using the varclus function in R.
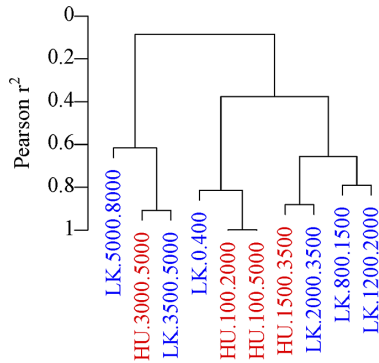
Figure 2. *Hierarchical clustering of two analyses of Edge scores based on the frequency bandings in (14), using Pearson correlations.*
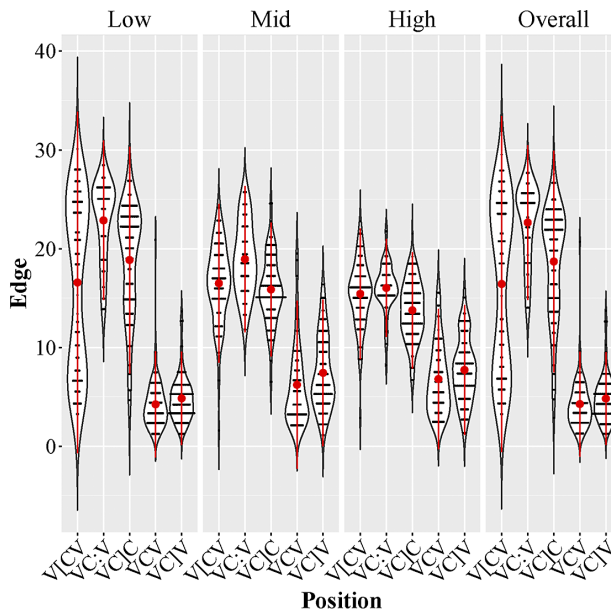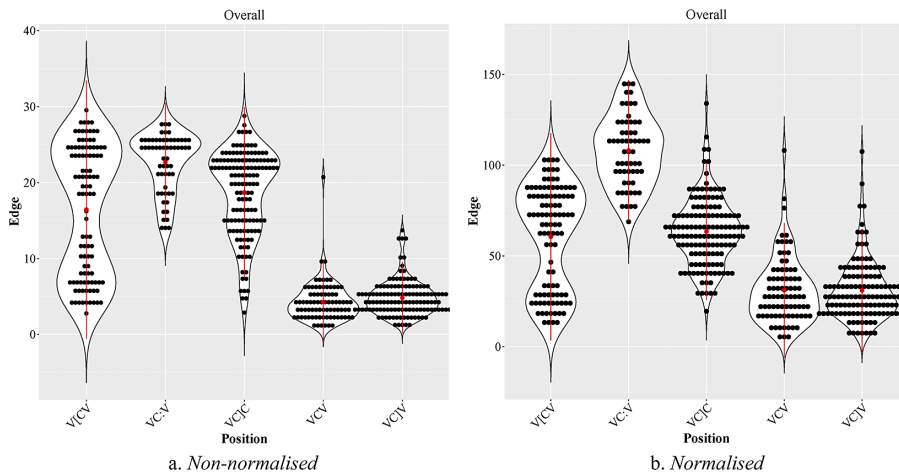

Figure 3. *Mean Edge scores by five stem positions and four frequency bands (four speakers). Black dots are individual data points; violins show distributions; red dots show means; red lines show standard deviations.*

A question arises as to whether the Edge value should be divided by the duration of the frame. Since normalising by duration is usual practice in phonetic studies using standard deviation measures (e.g., in investigations of pitch variation), we follow it in the statistical analyses to be presented below. However, we are not entirely confident that this is the right way to go with lenition, since it potentially prejudices our interpretation of the Edge results. It downplays the role of duration in lenition, placing the focus firmly on the impact the process has on articulatory stricture. It therefore implicitly subscribes to the view that lenition rests on a unidirectional causal link between duration and articulatory undershoot: Temporally compressing a consonant denies the speaker sufficient time to complete a planned articulatory closure (see, e.g., Cohen Priva 2017).

While the cause of lenition is not the subject of this article, we should bear in mind that there are good reasons to be sceptical of the undershoot account (summarised below in §6). Under an alternative interpretation, the duration and stricture dimensions of lenition are two correlates of an overall package

Figure 4. Comparison of raw and normalised Edge scores in the overall frequency band for five stem positions.

of planned speaker behaviour, with no commitment to whether one is causally linked to the other. That view argues against normalising the Edge results by duration.

Delving deeper into the issue of how duration and stricture are related in lenition would take us beyond the scope of this article. We limit ourselves here to a brief illustration of how the Edge results differ according to whether they are normalised by duration or not. This is shown in Figure 4, where we can compare the raw and normalised results for Edge in the overall frequency band.[13] Both methods clearly group VCV and VC]V together on relatively low Edge scores, confirming these positions as favouring vocalisation. The main difference lies in the place of geminates (VC:V). Without normalisation, geminates group closely with V[CV and VC]C on relatively high Edge scores, indicating that these positions resist vocalisation to roughly the same extent. With normalisation, all three of these positions also show up as resisting vocalisation, but with geminates doing so to a much greater extent.

To verify these initial observations, we conducted a Bayesian linear mixed effects analysis which compared the Edge scores pairwise across the five stem positions by constructing ten separate models for each of the four frequency bands.[14] Each of the models included a fixed effect for stem position and three random intercepts for speaker, word type and place of articulation. The last of these was included because it has been shown in other studies to have a significant influence on the intensity properties of consonants targeted by lenition (Lavoie 2001; Figueroa Candia 2016). To determine whether there is statistical evidence of a stem-position effect, a series of nested-model analyses was performed in which a null model omitting stem position as the fixed effect was compared with a full model including stem position. The level of evidence was measured using the change in LOOIC (Leave-One-Out Information Criterion) and the change in WAIC (Watanabe–Akaike Information Criterion) (see Vehtari *et al.* 2017). Substantial support for a stem effect can be inferred when a given analysis produces five or more changes in information criteria – a relatively conservative threshold.[15] A statistically well-supported

---

[13] The denominator of the standard deviation formula is the square root of the number of intervals. Normalising by duration mitigates the effect whereby the longer the frame is, the lower the standard deviation will be.

[14] We ran the Bayesian mixed effects linear regression in R (R Core Team 2017), using the brm function in the brms library (Bürkner 2017). For Edge (and for the Noise results reported below), we also ran two other types of statistical analyses. One was a version of the brm regression analysis, but with only one model fitted per frequency band. Another was *k*-means clustering. In both cases, the results were broadly in line with those of the analyses presented here.

[15] The changes in information criteria can be interpreted according to the following rules of thumb (see Raftery 1995; Burnham & Anderson 2002, 2004): Change in 1–2 units suggests little to no support, change in 4–7 units suggests substantial support and change in >10 units suggests full support. A threshold of 5 was chosen for this study.

result is one where the majority of the nested-model comparisons cross this threshold, indicating that the two stem positions being compared belong to different clusters. It is the statistically weakly supported results that we are primarily interested in, since they indicate where pairs of positions belong to the same cluster.

The results are summarised in Table 1. The sub-tables, showing the four different frequency bands, are organised as four-by-four grids to show the pairwise comparisons between stem positions. Cells containing results with only weak statistical support are highlighted. Of the four frequency bands, two come closest to providing statistical support for the expected partition of positions into lenition-resistant V[CV, VC:V, VC]C *vs.* lenition-prone VCV, VC]V. Both the overall band and the low band fail to distinguish significantly between members of the following pairs of stem positions: VC:V and V[CV; VC]C and V[CV; VCV and VC]V. The other two bands, mid and high, provide only partial support for the expected partition. On the one hand, they do not distinguish significantly between VC]C and V[CV or between VCV and VC]V. On the other, however, they also throw up unexpected positional pairings. Both fail to distinguish VC]C from VC]V. The high band also fails to distinguish between members of the following pairs: VCV and V[CV; VC]V and V[CV); VC]C and VCV. The analysis presented in Table 1 thus tells us that the main amplitude/intensity landmarks of vocalisation/spirantisation are best sought in the low and overall frequency bands.

5.3. *Aperiodic energy: 'Noise'*

5.3.1. *Noise: method*

The Edge measure tells us about how lenition shapes the way a consonant modulates the carrier signal along the dimension of amplitude. As we know from previous studies, this is the dimension primarily affected by what is traditionally called spirantisation. What the Edge measure does not tell us, however, is whether this term is accurate in characterising the resulting consonant as a canonical spirant/fricative. As we also know from previous studies, the process is better termed vocalisation in languages where it produces frictionless continuants. To answer this question, we need a different measure, one that gets a handle on the effect lenition has on periodicity. Ideally, such a measure should also be able to tell us something about the release properties of the final stops that alternate with vocalised outputs. As summarised in (8), Ibibio has a full series of these stops, which according to impressionistic descriptions are variably unreleased. This suggests they may lack the aperiodic burst that accompanies the release of stops in other positions.

One way of measuring the degree of periodicity in a sound is the harmonics-to-noise ratio, employed by Broś *et al.* (2021) in the analysis of Spanish lenition. The higher the ratio (measured in decibels), the more harmonic and thus more vowel-like the sound is. Broś *et al.* (2021) use this measure to gauge the extent to which so-called spirantisation produces approximant outcomes. The measure would be applicable to the analysis of the similar effect lenition has on (non-initial) intervocalic consonants in Ibibio. However, for Ibibio, we require a different measure, one that also allows us to investigate the release properties of final stops – which have no equivalent in Spanish.

The 'Noise' measure we employ here, adopted from Harris & Urua (2001), gauges the degree of aperiodic energy associated with a consonant by using an algorithm that combines separate measurements of amplitude and aperiodicity within the VCX analysis frame. First, the algorithm computes the aperiodicity of the signal as a function of time by using an autocorrelation measure to estimate what proportion of the signal is predictable from one moment to the next. Second, the algorithm locates the target consonant by finding the point in the analysis frame with the minimum amplitude. Third, it then searches forward within the frame until it finds the point with the maximum product of aperiodicity and normalised amplitude values, which yields a Noise score.[16] (Since the aperiodicity and amplitude

---

[16]The procedure for normalising amplitude expresses the energy of all points within the analysis frame as a simple fraction of the point with the greatest energy. For the details, see the `noisanal` function in the SFS manual (https://www.phon.ucl.ac.uk/resource/sfs/help/man/noisanal.htm).

Table 1. *Nested model comparisons of 10 Bayesian mixed effects regression models for Edge scores in each of four frequency bands.*

Band			VC:V	VC]C	VCV	VC]V
Overall	V[CV	ΔLOOIC	−3.24	−1.01	−12.76	−12.29
		ΔWAIC	−0.78	−0.79	−8.93	−6.54
	VC:V	ΔLOOIC		−31.20	−35.58	−28.74
		ΔWAIC		−27.66	−27.77	−8.98
	VC]C	ΔLOOIC			−34.18	−36.19
		ΔWAIC			−31.26	−30.89
	VCV	ΔLOOIC				−0.05
		ΔWAIC				−0.11
Low	V[CV	ΔLOOIC	−2.88	−0.05	−12.57	−9.86
		ΔWAIC	−0.99	−0.09	−8.85	−6.29
	VC:V	ΔLOOIC		−30.57	−35.36	−27.30
		ΔWAIC		−27.46	−27.43	−7.54
	VC]C	ΔLOOIC			−32.97	−36.61
		ΔWAIC			−30.46	−31.35
	VCV	ΔLOOIC				−0.53
		ΔWAIC				−0.44
Mid	V[CV	ΔLOOIC	−8.59	−0.32	−9.02	−7.71
		ΔWAIC	−5.35	−0.50	−8.56	−5.80
	VC:V	ΔLOOIC		−24.55	−15.91	−18.94
		ΔWAIC		−22.52	−13.37	−11.54
	VC]C	ΔLOOIC			−6.18	−0.52
		ΔWAIC			−5.63	−1.61
	VCV	ΔLOOIC				−1.11
		ΔWAIC				−1.11
High	V[CV	ΔLOOIC	−14.37	−1.68	−2.58	−2.06
		ΔWAIC	−11.89	−1.43	−2.72	−1.57
	VC:V	ΔLOOIC		−27.04	−12.96	−10.93
		ΔWAIC		−25.98	−11.55	−6.59
	VC]C	ΔLOOIC			−0.62	−0.18
		ΔWAIC			−0.59	−0.44
	VCV	ΔLOOIC				−0.88
		ΔWAIC				−0.94

factors are both normalised, the Noise score is itself also normalised, i.e., it ranges from 0 to 1.) The higher the Noise score, the greater the degree of aperiodic energy in the signal.

The Noise measure taps into the two main sources of aperiodic energy in speech: turbulent airflow during fricative articulations and transient bursts produced by the release of oral stops (Rosen 1992). It provides us with a way of determining the degree of aperiodicity in consonants occupying lenition-prone positions. In stem-internal intervocalic position (VCV), it indicates the extent to which spirantisation produces vocalic outputs (lower Noise scores) rather than canonical fricatives (higher Noise scores). In stem-final preconsonantal position (VC]C), it indicates the extent to which a stop is accompanied by a release burst: The more salient the burst, the higher the Noise score.
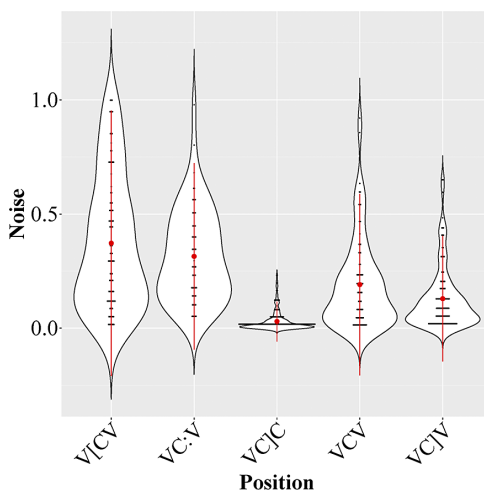
Figure 5. *Mean Noise values (aperiodicity × amplitude, normalised) by five stem positions (all four speakers).*

5.3.2. Noise: results

Figure 5 shows the distribution of Noise values across the five stem positions, averaged across all four speakers. The means for the two positions that resist lenition – stem-initial V[CV and geminate VC:V – are noticeably higher than those for the two positions that undergo spirantisation – non-initial intervocalic VCV and VC]V. The relatively high values of the former suggest that the Noise measure successfully captures the aperiodic energy of the release burst in unlenited plosives. The lower means of the spirantising positions suggest consonants with rather less aperiodic energy than would be expected of canonical fricatives. This echoes Spanish and reinforces the conclusion that this type of lenition is better described as vocalisation rather than the more traditional term spirantisation.

One position clearly stands apart from these four: stem-final preconsonantal VC]C, with a Noise mean close to zero. This position scores high on the Edge measure (see Figure 3), confirming it as hosting stops that resist spirantisation. Its very low Noise score indicates an almost complete absence of a release burst. The position also shows much less variation than the others, at least with the values conflated across all four speakers, as in Figure 5. (When the values are broken down by speaker, as shown in Figure A2 in Appendix A, we see a certain amount of variation between individuals. Speaker M1 in Figure A2 shows higher Noise values for position VC]C than the other speakers, indicating that he at least sometimes produces final stops with a release burst. This is consistent with phonetic descriptions of these consonants as being 'variably' released (Connell 1991; Urua 2000).)

To verify these initial observations, we followed a Bayesian linear mixed effects modelling procedure parallel to that used in the Edge analysis. We compared the Noise scores pairwise across the five stem positions by constructing ten separate models. Nested model comparisons were performed to determine if there is statistical evidence of a stem-position effect in each model. The results are summarised in Table 2. As before, our main interest lies in comparisons with statistically weakly supported results based on information criteria, since they indicate where pairs of positions belong to the same cluster.

According to this analysis, Noise fails to distinguish significantly between members of the following three pairs of stem positions: V[CV and VC:V; VCV and VC]V; V[CV and VCV. The first two of these groupings match the initial observations based on Figure 5. What is less expected is the pairing of V[CV with VCV. However, the difference between these two positions is the strongest of the three pairs, though still statistically weak (ΔLOOIC = 4.92, ΔWAIC = 2.50, compared to ΔLOOIC and ΔWAIC< 0.5 for the other two pairs). This is likely to reflect the wide spread of VCV values seen in Figure 5.

Table 2. *Nested model comparisons of 10 Bayesian mixed effects regression models for Noise scores.*

		VC:V	VC]C	VCV	VC]V
V[CV	ΔLOOIC	−0.11	−14.49	−4.92	−14.50
	ΔWAIC	−0.25	−5.86	−2.50	−10.10
VC:V	ΔLOOIC		−28.75	−7.33	−17.40
	ΔWAIC		−24.43	−6.17	−15.27
VC]C	ΔLOOIC			−23.44	−28.85
	ΔWAIC			−21.07	−23.92
VCV	ΔLOOIC				−0.08
	ΔWAIC				−0.04

5.4. Edge and Noise: discussion

The Edge and Noise analyses provide intersecting classifications of the Ibibio data by stem position, the separation being more clear-cut for certain speakers and positions than others. The positional classifications produced by the two analyses are different; this is unsurprising in light of the fact that they measure different acoustic dimensions.

The Edge analysis broadly divides the positions along the following lines: (a) stem-initial (V[CV), geminate (VC:V) and stem-final preconsonantal (VC]C) *vs.* (b) non-initial prevocalic (VCV and VC]V). The general directionality of this division is broadly in line with what the impressionistic descriptions given in §4 would lead us to expect. Consonants in Edge group (a) show relatively high levels of energy fluctuation across the VCX analysis frame, indicating resistance to lenition. Consonants in Edge group (b) have much lower Edge values, indicating a relatively shallow dip in energy during the transition between the flanking sonorants. This bears out the description of intervocalic stem-tail consonants as targets of spirantisation or vocalisation.

The Noise analysis yields a different classification of positions: (a) stem-initial (V[CV) and geminate (VC:V); (b) non-initial prevocalic singletons (VCV and VC]V) and (c) final preconsonantal (VC]C). Consonants in Noise group (a) show relatively high levels of aperiodic energy. Since the Edge analysis confirms these same two positions as resistant to vocalisation/spirantisation, the aperiodic energy here can reasonably be attributed to the noise burst accompanying plosive release. Consonants in Noise group (b), in contrast, show significantly lower levels of aperiodic energy, indicating a lack of continuous frication noise in lenited continuants. The very low Noise scores in group (c) indicate a preferred lack of burst release in final preconsonantal stops.

The intersection of Edge and Noise identifies stem-final preconsonantal (VC]C) as an intermediate position. Under the Edge analysis, this position patterns with V[CV in showing a marked drop in energy, indicative of a maintenance of stop closure during the consonant. Under the Noise analysis, on the other hand, it stands out on its own, showing a lower level of aperiodic energy than any of the other positions.

These results support the conclusion that the Ibibio consonants we have been examining come in three different modulation sizes, ranked from greatest to smallest: plosives (geminates and stem-initial singletons), unreleased stops (stem-final preconsonantal singletons) and continuants (non-initial prevocalic singletons). We will now attempt to provide an explicit measure of this ranking.

5.5. Modelling modulation size

The Edge and Noise analyses separately measure two of the acoustic dimensions along which lenition affects the way consonants modulate the carrier signal. As long as they remain separate measures, they

cannot be said to underpin a unified characterisation of lenition as an overall reduction in modulation size. In fact, it might be said that, in measuring amplitude, Edge on its own simply recapitulates the sonority-based account of lenition we critiqued earlier, and all we have done is add another dimension to the definition.

In response, we attempt to integrate our Edge and Noise measures by drawing on Ohala & Kawasaki-Fukumori's (Ohala & Kawasaki-Fukumori 1997) characterisation of the magnitude of a modulation as the distance it travels through an acoustic space defined by various parameters (see again the schematisation in Figure 1). The two-dimensional space defined by Edge and Noise cannot simply be constructed by linearly plotting one set of values against the other, because we know not all acoustic parameters are equally important to listeners. In other words, our integration method needs to mimic perceptual cue weighting. To achieve this, we first estimate the relative weighting of Edge and Noise across the different stem positions. We then visualise the modulations associated with the positions in a multidimensional, perceptually weighted acoustic space.

We approach the task by modelling how Ibibio listeners might distinguish different stem positions by weighting and integrating auditory–acoustic cues provided by Edge and Noise. The particular model we employ is Naive Discriminative Learning (NDL), incorporating the Rescorla–Wagner learning rule (Rescorla & Wagner 1972).[17] This has been shown to provide a psychologically plausible model of human learning in areas including lexical processing (Milin *et al.* 2017), morphological processing (Baayen *et al.* 2013), phonological concept learning (Moreton *et al.* 2017) and speech recognition (Arnold *et al.* 2017; see also Chen *et al.* 2007).

The model was trained on seven factors to predict the five stem positions: Noise, the four frequency bands of Edge, place of articulation and speaker. The four Edge bands are strongly correlated and thus decompose into four principal components.[18] To estimate cue weighting, we applied a non-parametric test to evaluate the relative importance of each predictor. As shown in Figure 6, the two most important variables for identifying stem positions are Noise and the first principal component of Edge. The relative importance of each of the acoustic predictors can be taken as the potential perceptual weight borne by that particular dimension. In principle, this would allow us to characterise the relative size of modulations in terms of the paths they trace through a multidimensional acoustic space, weighted by the dimensions' perceptual importance. The relevant dimensions are Noise and the four Edge parameters. However, because of the impracticality of visualising a five-dimensional space, we limit ourselves to the two most important predictors: Noise and the overall frequency component of Edge (the most representative component of Edge, since it covers all four of the bands).

For the purposes of integration, the values on both dimensions need to be normalised. Our Noise values are already in this format. The Edge values we have been working with need to be normalised because they are expressed as standard deviations (of mean amplitude) and as such have no upper limit (as least mathematically). A convenient way of achieving this is simply to divide each Edge value by the maximum Edge value found in the study.

We weight the two acoustic dimensions according to the variable importance values provided by the NDL model in Figure 6. To do this, we take the greatest importance value from amongst the Edge parameters (0.0634) and divide it by the importance value of Noise (0.0793), yielding a ratio of 0.7995, by which we then divide the Noise value.

---

[17]We implemented the Rescorla–Wagner method using the Naive Discriminative Learning library by Arppe *et al.* (2015) in R. This implementation has the advantage that the discriminative learning is completely driven by the input we provide. The parameters of the Rescorla–Wagner equations are derived from the equilibrium equations of Danks (2003), as a consequence of which there are no free parameters we could manipulate.

[18]For Edge, PCA 1 is the major principal component, with 80.65% of variance. Edge PCA 2, PCA 3 and PCA 4 have 16.62%, 2.71% and 0.017% of variance, respectively. Edge PCA 1 corresponds to Edge frequency-bands Low and Overall (with correlation values (Rs) of 0.95) and also partly to Mid (R:0.86) and High (R:0.76). Edge PCA 2 corresponds mostly to High (R:0.60), followed by Mid (R:0.45); it correlates positively with Low (R: 0.31) and Overall (R: 0.30). Edge PCA 3 corresponds to High (R:0.26) and Mid (R: 0.25); it does not correspond to Low and Overall (|R|s < 0.02). Edge PCA 4 does not correlate with any of the bands (all |R|s < 0.02).
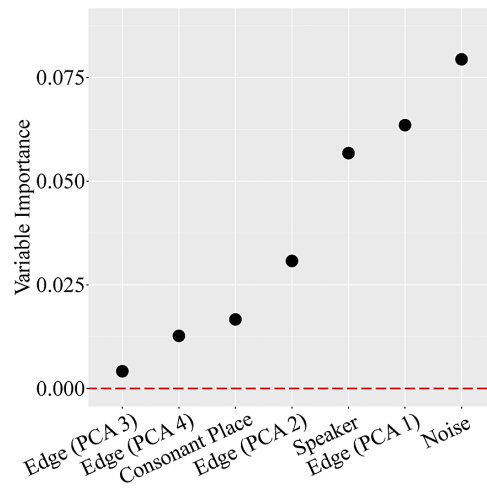
Figure 6. *Relative importance of seven variables in predicting Ibibio stem positions, generated by the NDL model described in the text. Any variable with a value above the dashed red line is a significant predictor.*
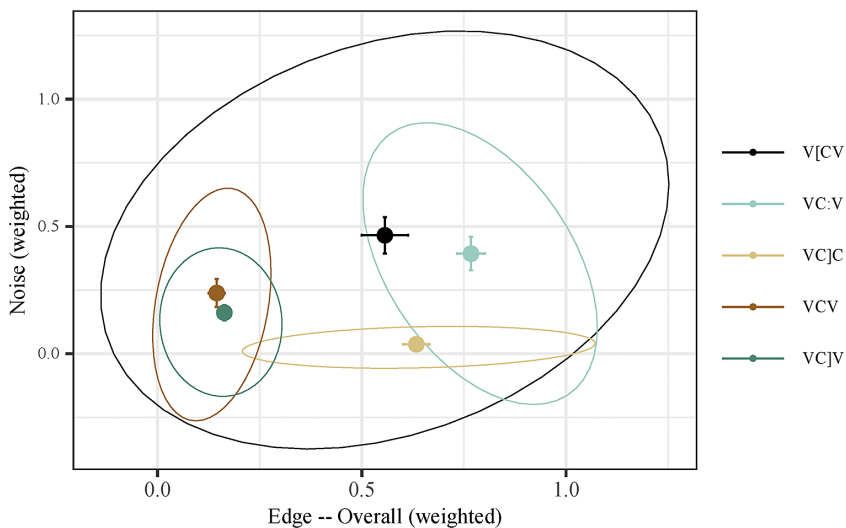
Figure 7. *Visualisation of the extent to which Ibibio consonants modulate the carrier signal (all speakers).*

Figure 7 plots the weighted Noise values (*y*-axis) against the weighted Edge values (*x*-axis), averaged across all speakers. Individual tokens of each stem position are grouped using ellipses; mean and error bars (95% confidence interval) are shown as points and crosses, respectively. The farther a token lies towards the top right of the plot, the greater the modulation. The resulting plot can be viewed as the perceptual space occupied by the five different stem positions.

Figure 7 allows us to visualise the extent to which the consonant in each word token modulates the carrier. The further a token lies towards the top right of the plot, the greater the modulation. We can see how unlenited stem-initial singleton consonants and medial geminates modulate the carrier to a greater extent than lenited non-initial prevocalic singletons. The picture also captures the intermediate size of the modulations produced by stem-final preconsonantal stops.

6. Conclusion

Consonants vary in the extent to which they modulate the carrier signal in speech. This variation lies at the heart of lenition, which can be understood as a unitary phenomenon that diminishes a consonant's impact on the carrier. We have shown here how this notion can be quantified. When lenition is positionally sensitive, as it is in Ibibio, we are presented with a moving picture of speech in which consonant modulations differ in magnitude from one moment to the next.

There is an intuitive appeal in having differences in modulation size encoded in phonology in a relatively direct way. The design properties of an MC-inspired feature framework that makes this possible have been outlined elsewhere (Harris & Lindsey 1995; Harris 2004, 2009). Every feature is monovalent and is phonetically defined in terms of how it modulates the carrier signal. The greater the degree to which a segment modulates the carrier, the more features it bears. As a result, segments in positions that resist lenition bear richer feature specifications than those that succumb to it. This unequal distribution of feature capital can be understood as part of a more general scheme involving positional strength or prominence. It aligns segmental phonology with prosodic properties such as stress or pitch accent. In a stress system, a strong syllable bears a certain specification – an accent mark – that weak syllables lack. In a system with vocalising lenition, a strong consonant bears a specification – an abruptness or 'edge' feature, say – that is missing from weak consonants. A tone language such as Ibibio illustrates how segmental asymmetries can signal accentual prominence even in the absence of stress (cf. Hyman 2008; Downing 2010; Harris & Hyman 2022).

This brings us back to the causation question touched on near the beginning of this article: why do grammars have lenition in the first place? In presenting the MC account, we have said nothing about the traditional assumption that lenition is driven by ease of articulation. Could modulation reduction be motivated by a pressure on speakers to minimise articulatory effort? There are at least two reasons for not following this line of enquiry.

The first is that the effort-based account is known to be problematic in and of itself. The amount of articulatory energy saved by producing a lenited version of a target consonant is in all likelihood too miniscule to warrant serious consideration as a driver of sound change (Kingston 2008). In any case, the account makes wrong predictions about the phonological conditions lenition should be subject to when it establishes itself in grammars. For example, we would be led to expect, wrongly, that opening types of lenition such as vocalisation and spirantisation should be sensitive to the relative aperture of the surrounding vowels (Kingston 2008).

The second reason for not pursuing the ease-of-articulation account is that there is a growing body of evidence favouring a more thoroughly listener-centred approach to lenition (e.g., Harris 2004; Kingston 2008; Shiraishi 2009; Kaplan 2011; Katz 2016; Katz & Fricke 2018). Leniting consonants in weak positions accentuates the auditory–acoustic contrast with unlenited consonants in strong positions. Moreover, it does so in a way that concentrates perceptually more salient consonants – those that modulate the carrier more – in strong positions. Various researchers have made this point in relation to intensity (Kingston 2008; Kaplan 2011; Katz 2016; Katz & Fricke 2018). From an MC perspective, the contrast is broader than this, encompassing not just intensity but also the other acoustic dimensions along which the carrier signal is modulated. Given the nature of lenition in Ibibio, we have focused here on intensity/amplitude and periodicity. By hypothesis, the most relevant dimension to study in a system with debuccalisation is spectral shape.

When the auditory–acoustic asymmetry between lenition-prone and lenition-resistant positions is harnessed to prosodic or morphosyntactic domain structure, it has the potential to provide parsing cues that aid the listener in segmenting the speech stream. This is the scenario in Ibibio, summarised in (8), where consonants resist lenition when in the head position of the stem and fall prey to it only when in the tail. We thus come to understand lenition as serving a demarcative function analogous to that of fixed word stress (cf. Trubetzkoy 1939; Hyman 1977; Cutler & Norris 1988). This suggests that the phonetic effects of lenition mesh with the range of segmental properties that have been found to provide listeners

with parsing cues, including aspiration, glottalisation and (non-)release in English stops (e.g., Abramson & Lisker 1970; Christie Jr. 1974; Church 1987; Jusczyk *et al.* 1999; Mattys & Melhorn 2007).

The neatest demarcative picture is where consonants only undergo lenition in positions that are internal or final within some domain, such as the word (Harris 2003; Katz & Fricke 2018; Katz & Pitzanti 2019; White *et al.* 2020) or phrase (Kingston 2008; Shiraishi 2009). Under these circumstances, consonants that resist lenition (or undergo fortition) help mark out the beginning of a domain.

The picture is perhaps not quite so neat when consonants are non-initial but nevertheless resist lenition because they occur in a locally 'protected' environment (Lass & Anderson 1975; Scheer & Ségéral 2008). This is the situation in Spanish, for example, where a nasal blocks spirantisation of a following voiced stop (see Hualde *et al.* 2011 for discussion and references); as a result, a strong consonant is not a totally reliable cue to domain-initial position in Spanish. A rather similar situation obtains in Ibibio, where domain-internal geminates remain inalterable. However, in this case, the protected consonants only occur in specific verbal paradigms and are part of a broader scenario where the relative strength of consonants distinguishes different positions within the stem. Alongside other cues to stem structure, including vowel harmony and tonal alternations, internal geminates thus help signal particular morphological categories. More generally, this means we can think of lenition as having the functional potential not just to mark out domains but also to aid morphosyntactic labelling.

This is not to write the speaker out of the lenition story altogether. The listener-centred account depicts lenition as fitting into a collaborative communicative enterprise involving both listener and speaker. In speech production, weak phonological positions are characterised by less extreme (hypoarticulated) gestures and strong positions by more extreme (hyperarticulated) gestures (see, e.g., Lindblom 1990; Pierrehumbert & Talkin 1992; Fougeron & Keating 1997; de Jong 1998; Keating *et al.* 2004). Lenition can be thought of as a phonologically entrenched reflex of hypoarticulation. However, rather than being driven by some speaker-centred notion of economy of effort, hypoarticulation in this scenario is part of planned speech behaviour that benefits listeners (cf. Xu & Prom-on 2019). Speakers execute more extreme gestures in order to produce greater modulations of the carrier in order to direct
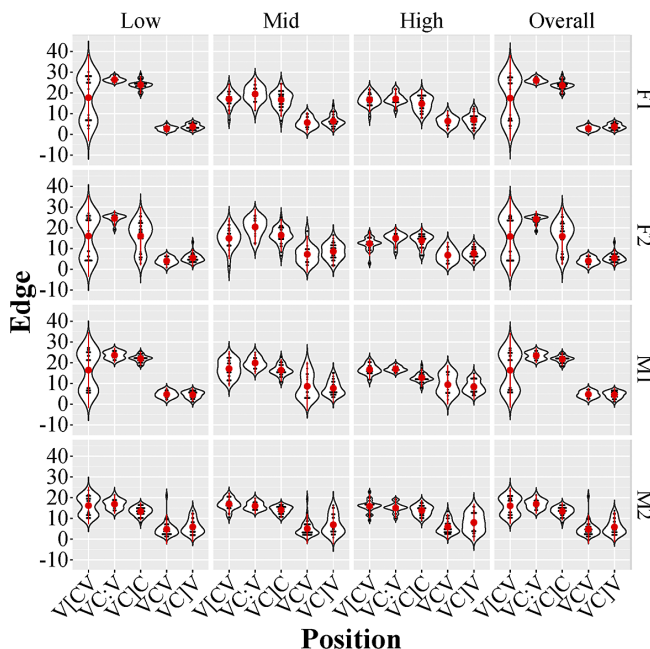
Figure A1. *Mean Edge scores (SD of amplitude in dB) by five stem positions, four frequency bands and four speakers (F1, F2, M1, M2).*
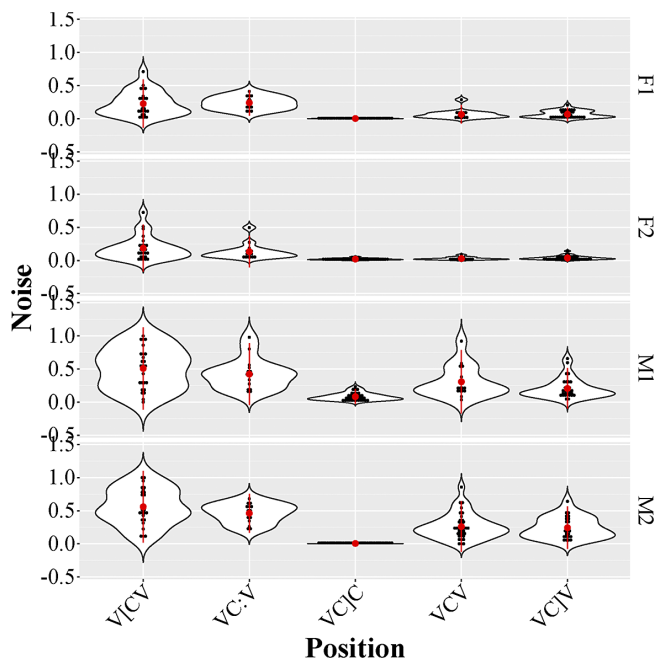
Figure A2. *Noise values (aperiodicty × amplitude) by five stem positions and four speakers.*

listeners' attention towards strong phonological positions. Hyperarticulation is thereby matched by 'hyperperception' in strong positions and hypoarticulation by 'hypoperception' in weak positions (de Jong 2000).

A. Additional figures

The figures in this appendix present speaker-by-speaker breakdowns of results that were presented in aggregated form in the main text. Figure A1 does this for the Edge scores discussed in §5.2.2 (cf. Figure 3).

Figure A2 shows speaker-by-speaker results for Noise scores (cf. Figure 5 in §5.3.2).

Competing interests. The authors declare no competing interests.

References

Abramson, Arthur S. & Leigh Lisker (1970). Discriminability along the voicing continuum: cross-language tests. In *Proceedings of the sixth International Congress of Phonetic Sciences*. Prague: Academia Publishing House of the Czechoslovak Academy of Sciences, 569–573.

Akinlabi, Akinbiyi & Eno E. Urua (1992). Prosodic target and vocalic specification in the Ibibio verb. *WCCFL* **11**, 1–14.

Akinlabi, Akinbiyi & Eno E. Urua (1994). Foot in the Ibibio verb. Ms., Rutgers University and University of Uyo.

Albert, Aviad & Bruno Nicenboim (2022). Modeling sonority in terms of pitch intelligibility with the nucleus attraction principle. *Cognitive Science* **46**, Article no. e13161, 68 pp.

Arnold, Denis, Fabian Tomaschek, Konstantin Sering, Florence Lopez & R. Harald Baayen (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates

between meanings straight from smart acoustic features, bypassing the phoneme as recognition unit. *PLoS One* **12**, Article no. e0174623, 16 pp.

Arppe, Antti, Peter Hendrix, Petar Milin, R. Harald Baayen, Tino Sering & Cyrus Shaoul (2015). ndl: naïve discriminative learning. R package. Version 0.2.17. https://CRAN.R-project.org/package=ndl.

Avery, Peter & William J. Idsardi (2001). Laryngeal dimensions, completion and enhancement. In T. Alan Hall (ed.) *Distinctive feature theory*, number 2 in Phonology and Phonetics. Berlin: Mouton de Gruyter, 41–70.

Baayen, R. Harald, Anna Endresen, Laura A. Janda, Anastasia Makarova & Tore Nesset (2013). Making choices in Russian: pros and cons of statistical methods for rival forms. *Russian Linguistics* **37**, 253–291.

Barnes, Jonathan (2006). *Strength and weakness at the interface: positional neutralization in phonetics and phonology*, number 10 in Phonology and Phonetics. Berlin: Mouton de Gruyter.

Bauer, Laurie (2008). Lenition revisited. *Journal of Linguistics* **44**, 605–624.

Beckman, Jill N. (1999). *Positional faithfulness: an Optimality Theoretic treatment of phonological asymmetries*. New York: Routledge.

Blevins, Juliette (2004). *Evolutionary phonology*. Cambridge: Cambridge University Press.

Botma, Bert & Marijn van 't Veer (2013). A fraction too much friction: the phonological status of voiced fricatives. In Suzanne Aalberse & Anita Auer (eds.) *Linguistics in the Netherlands 2013*. Amsterdam: John Benjamins, 46–60.

Broś, Karolina, Marzena Żygis, Adam Sikorski & Jan Wołłejko (2021). Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. *Phonology* **38**, 1–40.

Bürkner, Paul-Christian (2017). brms: an R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1–28.

Burnham, Kenneth P. & David R. Anderson (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer.

Burnham, Kenneth P. & David R. Anderson (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* **33**, 261–304.

Carrasco, Patricio, José I. Hualde & Miquel Simonet (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica* **69**, 149–179.

Chen, Matthew Y. (1976). From Middle Chinese to Modern Pekin. *Journal of Chinese Linguistics* **4**, 113–277.

Chen, Zhe, Simon Haykin, Jos J. Eggermont & Suzanna Becker (2007). *Correlative learning: a basis for brain and adaptive systems*. Hoboken: Wiley-Interscience.

Christie, Jr., William M. (1974). Some cues for syllable juncture perception in English. *JASA* **55**, 819–821.

Church, Kenneth W. (1987). Phonological parsing and lexical retrieval. *Cognition* **25**, 53–69.

Clements, G. N. (2009). Does sonority have a phonetic basis? In Eric Raimy & Charles E. Cairns (eds.) *Contemporary views on architecture and representations in phonology*, number 48 in Current Studies in Linguistics. Cambridge, MA: MIT Press, 165–175.

Cohen Priva, Uriel (2017). Informativity and the actuation of lenition. *Lg* **93**, 569–597.

Colantoni, Laura & Irina Marinescu (2010). The scope of stop weakening in Argentine Spanish. In Marta Ortega-Llebaria (ed.) *Selected proceedings of the 4th conference on laboratory approaches to Spanish phonology*. Somerville, MA: Cascadilla Proceedings Project, 100–114.

Cole, Jennifer, José I. Hualde & Khalil Iskarous (1998). Effects of prosodic and segmental context on /g/-lenition in Spanish. In Osamu Fujimura, Brian D. Joseph & Bohumil Palek (eds.) *Proceedings of LP98*. Prague: Charles University Press (Karolinum), 575–589.

Coleman, John, John Aston & Davide Pigole (2015). Reconstructing the sounds of words from the past. In Maria Wolters, Judy Livingstone, Bernie Beattie, Rachel Smith, Mike MacMahon, Jane Stuart-Smith & Jim Scobbie (eds.) *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow, 5 pp.

Connell, Bruce (1991). *Phonetic aspects of the lower cross languages and their implications for sound change*. PhD dissertation, University of Edinburgh.

Cook, Thomas Louis (1985). *An integrated phonology of Efik*. PhD dissertation, Universiteit Leiden.

Cutler, Anne & Dennis Norris (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception & Performance* **14**, 113–121.

Danks, David (2003). Equilibria of the Rescorla–Wagner model. *Journal of Mathematical Psychology* **47**, 109–121.

de Jong, Kenneth (1998). Stress-related variation in the articulation of coda alveolar stops: flapping revisited. *JPh* **26**, 283–310.

de Jong, Kenneth (2000). Attention modulation and the formal properties of stress systems. *CLS* **36**, 71–92.

Downing, Laura J. (2006). *Canonical forms in prosodic morphology*, number 12 in Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

Downing, Laura J. (2010). Accent in African languages. In Harry van der Hulst, Rob Goedemans & Ellen van Zanten (eds.) *A survey of word accentual patterns in the languages of the world*. Berlin: De Gruyter Mouton, 381–428.

Dudley, Homer (1940). The carrier nature of speech. *Bell Systems Technical Journal* **19**, 495–515.

Ennever, Thomas, Felicity Meakins & Erich R. Round (2017). A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology* **8**, Article no. 20, 32 pp.

Figueroa Candia, Mauricio Alejandro (2016). *Lenition in the production and perception of Chilean Spanish approximant consonants: implications for lexical access models*. PhD dissertation, University College London.

Flemming, Edward S. (1995). *Auditory representations in phonology*. PhD dissertation, University of California, Los Angeles.

Foley, James (1977). *Foundations of theoretical phonology*, number 20 in Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

Fougeron, Cécile & Patricia A. Keating (1997). Articulatory strengthening at edges of prosodic domains. *JASA* **101**, 3728–3740.

Gurevich, Naomi (2004). *Lenition and contrast: the functional consequences of certain phonologically conditioned sound changes*. New York: Routledge.

Gurevich, Naomi (2011). Lenition. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*, volume 3. Oxford: Wiley-Blackwell, 1559–1575.

Hankamer, Jorge & Judith Aissen (1974). Attention modulation and the formal properties of stress systems. *CLS* **10**, 131–145.

Harris, James W. (1969). *Spanish phonology*. Cambridge, MA: MIT Press.

Harris, John (1994). *English sound structure*. Oxford: Blackwell.

Harris, John (1998). Licensing inheritance: an integrated theory of neutralisation. *Phonology* **14**, 315–370.

Harris, John (2003). Grammar-internal and grammar-external assimilation. In Maria-Josep Solé, Daniel Recasens & Joaquin Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Adelaide: Causal Productions, 281–284.

Harris, John (2004). Release the captive coda: the foot as a domain of phonetic interpretation. In John Local, Richard Ogden & Rosalind Temple (eds.) *Phonetic interpretation*, number VI in Papers in Laboratory Phonology. Cambridge: Cambridge University Press, 103–129.

Harris, John (2005). Vowel reduction as information loss. In Philip Carr, Jacques Durand & Colin J. Ewen (eds.) *Headhood, elements, specification and contrastivity*, number 259 in Current Issues in Linguistic Theory. Amsterdam: John Benjamins, 119–132.

Harris, John (2006). The phonology of being understood: further arguments against sonority. *Lingua* **116**, 1483–1494.

Harris, John (2009). Why final devoicing is weakening. In Kuniya Nasukawa & Phillip Backley (eds.) *Strength relations in phonology*, number 103 in Studies in Generative Grammar. Berlin: Mouton de Gruyter, 9–46.

Harris, John (2013). Wide-domain r-effects in English. *JL* **49**, 329–365.

Harris, John & Larry M. Hyman (2022). Segmental prominence and the modulated carrier signal. In Moses Effiong Ekpenyong & Imelda Icheji Udoh (eds.) *Current issues in descriptive linguistics and digital humanities: a festschrift in honor of Professor Eno-Abasi Essien Urual*. Singapore: Springer, 487–499.

Harris, John & Geoff Lindsey (1995). The elements of phonological representation. In Jacques Durand & Francis Katamba (eds.) *Frontiers of phonology: atoms, structures, derivations*. Harlow: Longman, 34–79.

Harris, John & Eno-Abasi Urua (2001). Lenition degrades information: consonant allophony in Ibibio. *Speech, Hearing and Language: Work in Progress* **13**, 72–105.

Hayes, Bruce (1986). Inalterability in CV phonology. *Lg* **62**, 321–351.

Henke, Eric, Ellen M. Kaisse & Richard Wright (2012). Is the Sonority Sequencing Principle an epiphenomenon? In Steve Parker (ed.) *The sonority controversy*, number 18 in Phonology and Phonetics. Berlin: De Gruyter Mouton, 65–100.

Hock, Hans Henrich (1991). *Principles of historical linguistics*. Berlin: Mouton de Gruyter.

Honeybone, Patrick (2008). Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.) *Lenition and fortition*. Berlin: Walter de Gruyter, 9–92.

Hualde, J., Miquel Simonet, Ryan Shosted & Marianna Nadeu (2010). Quantifying Iberian spirantization: acoustics and articulation. In Julia Herschensohn (ed.) *Romance linguistics 2010: selected proceedings of the 40th Symposium on Romance Languages*. Amsterdam: John Benjamins, 26–28.

Hualde, José Ignacio, Miquel Simonet & Marianna Nadeu (2011). Consonant lenition and phonological recategorization. *Laboratory Phonology* **2**, 301–329.

Hyman, Larry M. (1975). *Phonology: theory and analysis*. New York: Holt, Rinehart & Winston.

Hyman, Larry M. (1977). On the nature of linguistic stress. *Southern California Occasional Papers in Linguistics* **4**, 37–82.

Hyman, Larry M. (1982). The representation of nasality in Gokana. In Harry van der Hulst & Norval Smith (eds.) *The structure of phonological representations (part 1)*. Dordrecht: Foris, 111–130.

Hyman, Larry M. (1990). Non-exhaustive syllabification: evidence from Nigeria and Cameroon. *CLS* **26**, 175–195.

Hyman, Larry M. (2003). Basaá (A.43). In Derek Nurse & Gérard Philippson (eds.) *The Bantu languages*. London: Routledge, 257–282.

Hyman, Larry M. (2008). Directional asymmetries in the morphology and phonology of words, with special reference to Bantu. *Linguistics* **46**, 309–350.

Hyman, Larry M., Nicholas Rolle, Hannah Sande, Emily Clem, Peter S. E. Jenks, Florian Lionnet, John Merrill & Nico Baier (2019). Niger–Congo linguistic features and typology. In H. Ekkehard Wolff (ed.) *The Cambridge handbook of African linguistics*. Cambridge: Cambridge University Press, 191–245.

Iverson, Gregory K. & Joseph C. Salmons (2003). Laryngeal enhancement in early Germanic. *Phonology* **20**, 43–74.

Jansen, Wouter (2004). *Laryngeal contrast and phonetic voicing: a laboratory phonology approach to English, Hungarian, and Dutch*. PhD dissertation, University of Groningen.

Jensen, John T. (2000). Against ambisyllabicity. *Phonology* **17**, 187–235.

Jusczyk, Peter W., Elizabeth A. Hohne & Angela Bauman (1999). Infants' sensitivity to allophonic cues for word segmentation. *Attention, Perception, & Psychophysics* **61**, 1465–1476.

Kaplan, Abby (2011). Perceptual pressures on lenition. *Language and Speech* **54**, 285–305.

Katz, Jonah (2016). Lenition, perception and neutralisation. *Phonology* **33**, 43–85.

Katz, Jonah & Melinda Fricke (2018). Auditory disruption improves word segmentation: a functional basis for lenition phenomena. *Glossa* **3**, Article no. 38, 25 pp.

Katz, Jonah & Gianmarco Pitzanti (2019). The phonetics and phonology of lenition: a Campidanese Sardinian case study. *Laboratory Phonology* **10**, Article no. 16, 40 pp.

Keating, Patricia, Taehong Cho, Cécile Fougeron & Chai-Shune Hsu (2004). Domain-initial strengthening in four languages. In John Local, Richard Ogden & Rosalind Temple (eds.) *Phonetic interpretation*, number 4 in Papers in Laboratory Phonology. Cambridge: Cambridge University Press, 145–163.

Kingston, John (2008). Lenition. In Laura Colantoni & Jeffrey Steele (eds.) *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. Somerville, MA: Cascadilla Proceedings Project, 1–31.

Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *LI* **10**, 421–441.

Kirchner, Robert (2004). Consonant lenition. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.) *Phonetically based phonology*. Cambridge: Cambridge University Press, 313–345.

Kirchner, Robert Martin (1998). *An effort-based approach to consonant lenition*. PhD dissertation, University of California, Los Angeles.

Ladefoged, Peter (2001). *A course in phonetics*. Fort Worth, TX: Harcourt.

Ladefoged, Peter & Ian Maddieson (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lass, Roger & John M. Anderson (1975). *Old English phonology*. Cambridge: Cambridge University Press.

Lavoie, Lisa M. (2001). *Consonant strength: phonological patterns and phonetic manifestation*. New York: Garland.

Lindblom, Björn (1990). Explaining phonetic variation: a sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer, 403–439.

Liu, Sharlene A. (1996). Landmark detection for distinctive feature-based speech recognition. *JASA* **100**, 3417–3430.

Mattys, Sven L. & James F. Melhorn (2007). Sentential, lexical, and acoustic effects on the perception of word boundaries. *JASA* **122**, 554–567.

Milin, Petar, Laurie Beth Feldman, Michael Ramscar, Peter Hendrix & R. Harald Baayen (2017). Discrimination in lexical decision. *PLoS One* **12**, Article no. e0171935, 42 pp.

Moreton, Elliott, Joe Pater & Katya Pertsova (2017). Phonological concept learning. *Cognitive Science* **41**, 4–69.

Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.) *The production of speech*. New York: Springer, 189–216.

Ohala, John J. & Haruko Kawasaki-Fukumori (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Stig Eliasson & Ernst Håkon Jahr (eds.) *Language and its ecology: essays in memory of Einar Haugen*. Berlin: Mouton de Gruyter, 343–365.

Ortega-Llebaria, Marta (2003). Effects of phonetic and inventory constraints in the spirantization of intervocalic voiced stops: comparing two different measures of energy change. In Maria-Josep Solé, Daniel Recasens & Joaquin Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Adelaide: Causal Productions, 2817–2820.

Parker, Stephen G. (2002). *Quantifying the sonority hierarchy*. PhD dissertation, University of Massachusetts Amherst.

Parker, Steve (2011). Sonority. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*, volume 2. Oxford: Wiley-Blackwell, 1160–1184.

Pierrehumbert, Janet & David Talkin (1992). Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (eds.) *Gesture, segment, prosody*, number 2 in Papers in Laboratory Phonology. Cambridge: Cambridge University Press, 90–116.

R Core Team (2017). *R: a language and environment for statistical computing [Computer program]*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org.

Raftery, Adrian E. (1995). Bayesian model selection in social research. *Sociological Methodology* **25**, 111–163.

Rescorla, Robert A. & Allan R. Wagner (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Abraham H. Black & William F. Prokasy (eds.) *Classical conditioning II: current research and theory*. New York: Appleton-Century-Crofts, 64–99.

Rice, Keren & Peter Avery (1989). On the interaction between sonorancy and voicing. *Toronto Working Papers in Linguistics* **10**, 65–82.

Romero Gallego, Joaquin (1995). *Gestural organization in Spanish: an experimental study of spirantization and aspiration*. PhD dissertation, University of Connecticut.

Rosen, Stuart (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **336**, 367–373.

Scheer, Tobias & Philippe Ségéral (2008). Positional factors in lenition and fortition. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.) *Lenition and fortition*. Berlin: Walter de Gruyter, 131–172.

Shiraishi, Hidetoshi (2009). Modelling initial weakenings. In Kuniya Nasukawa & Phillip Backley (eds.) *Strength relations in phonology*, number 103 in Studies in Generative Grammar. Berlin: Mouton de Gruyter, 183–218.

Smith, Jennifer L. (2008). Markedness, faithfulness, positions, and contexts: lenition and fortition in Optimality Theory. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.) *Lenition and fortition*. Berlin: Walter de Gruyter, 519–560.

Soler, Antonia & Joaquín Romero (1999). The role of duration in stop lenition in Spanish. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey (eds.) *Proceedings of the 14th International Congress of Phonetic Sciences*. Berkeley, CA: University of California, 483–486.

Stevens, Kenneth N. (1985). Evidence for the role of acoustic boundaries in the perception of speech sounds. In Victoria A. Fromkin (ed.) *Phonetic linguistics: essays in honor of Peter Ladefoged*. Orlando: Academic Press, 243–255.

Szigetvári, Péter (2008). What and where? In *Joaquim Brandão de Carvalho*, Tobias Scheer & Philippe Ségéral (eds.) *Lenition and fortition*. Berlin: Walter de Gruyter, 93–129.

Tang, Kevin, Ratree Wayland, Fenqi Wang, Sophia Vellozzi, Rahul Sengupta & Lori Altmann (2023). From sonority hierarchy to posterior probability as a measure of lenition: the case of Spanish stops. *JASA* **153**, 1191–1203.

Traunmüller, Hartmut (1994). Conventional, biological and environmental factors in speech communication: a modulation theory. *Phonetica* **51**, 170–183.

Traunmüller, Hartmut (2005). *Speech considered as modulated voice*. Ms, Stockholm University. Available at https://www.researchgate.net/profile/Hartmut-Traunmueller/publication/250390908_Speech_considered_as_modulated_voice/links/5bab9872a6fdccd3cb75c4ac/Speech-considered-as-modulated-voice.pdf.

Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie. Travaux du cercle linguistique de Prague*.

Trujillo, Ramón (1980). Sonorización de sordas en Canarias. *Anuario de Letras: Lingüística y Filología* **18**, 247–253.

Urua, Eno Essien (1990). *Aspects of Ibibio phonology and morphology*. PhD dissertation, University of Ibadan.

Urua, Eno-Abasi Essien (2000). *Ibibio phonetics and phonology*. Cape Town: Centre for Advanced Studies of African Society.

Vehtari, Aki, Andrew Gelman & Jonah Gabry (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432.

Westbury, John R. & Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *JL* **22**, 145–166.

White, Laurence, Silvia Benavides-Varela & Katalin Mády (2020). Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *JPh* **81**, Article no. 100982, 44 pp.

Wright, Richard (2001). Perceptual cues in contrast maintenance. In Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. San Diego: Academic Press, 251–277.

Wright, Richard (2004). A review of perceptual cues and robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.) *Phonetically based phonology*. Cambridge: Cambridge University Press, 34–57.

Xu, Yi & Santitham Prom-on (2019). Economy of effort or maximum rate of information? Exploring basic principles of articulatory dynamics. *Frontiers in Psychology* **10**, Article no. 2469, 22 pp.