INFORMATIVITY AND THE ACTUATION OF LENITION

Uriel Cohen Priva

Brown University

What causes Indonesian to lenite word-final /k/, American English to lenite word-final /t/, and Spanish to lenite word-final /s/? This article shows that all three observed lenition patterns can be motivated using a single principle: languages preferentially lenite segments that provide relatively low informativity compared to the amount of informativity those segments carry in other languages. In a comparison of a diverse sample of seven languages from the LDC CALLHOME and CALLFRIEND corpora, Indonesian /k/, American English /t/, and Spanish /s/ are found to have the lowest informativity, predicting that they would be more likely to be affected by sound change processes. In a subsequent regression-based corpus study, low informativity predicted the propensity of word-final lenition of all obstruents in American English after phonetic and phonological factors were controlled for. This article therefore provides a partial solution to the famous actuation problem (Weinreich et al. 1968) with respect to the actuation of lenition processes.*

Keywords: lenition, information, informativity, actuation problem, sound change

1. Introduction. Different languages lenite (voice, spirantize, approximate, debuccalize, or delete)[1] different sounds in word-final positions. American English variably lenites /t/ word-finally (Kahn 1976, Zue & Laferriere 1979), varieties of Spanish variably lenite /s/ word-finally (e.g. Poplack 1980, Hochberg 1986, Fox 2001), and Indonesian lenites /k/ word-finally (Soderberg & Olson 2008). English, Indonesian, and Spanish all allow /s/, /t/, and /k/ word-finally; why is it that each one preferentially lenites a different final segment? More generally, is it possible to predict which segments are most vulnerable to lenition within a given language? Answering this question can provide a partial answer to the famous actuation problem (Weinreich et al. 1968): what causes a potential sound change to occur in a particular language at a particular time?

Although sound change has gathered considerable attention in linguistic research (e.g. Hockett 1965, Ohala 1993a, Kiparsky 1995, Labov 2001, Hale 2003, Blevins 2004, Kingston 2007), until recently few accounts tried to tackle the actuation problem directly (see Stevens & Harrington 2014 for a review). Recent work identifies several aspects that affect the actuation of sound change, focusing on the contribution of individuals to this process. Baker and colleagues (2011), Garrett and Johnson (2013), and Yu (2013) studied the kinds of variation in sounds that might make them susceptible to spread by individuals, and additionally investigated which individuals might be more likely to be responsible for the initiation and spread of sound change.

Following structuralist tradition (Hockett 1955), Cohen Priva 2012 and Wedel, Kaplan, and Jackson (2013) point to quantifiable preservation forces as inhibitors of change. They argue that change results from a balance between phonetic pressures that make change likely and communicational pressures that forbid some kinds of change to occur. This approach has been recently modeled in Sóskuthy 2015, which demonstrates computationally that communicative pressures play a significant role in keeping linguistic systems stable.

This article also relies on information-preservation principles in order to account for the distribution of word-final lenition processes. I argue that it is possible to predict which language is likely to lenite which segment by quantifying its language-specific

* I would like to thank Dan Jurafsky, Paul Kiparsky, Emily Gleason, Andy Wedel, Meghan Sumner, Matthew Adams, Arto Anttila, Vera Gribanova, Roey Gafter, and Pauline Jacobson for many helpful discussions and their great assistance in the different stages and incarnations of this article.

[1] See §2.2 for the types of lenition processes this article aims to address.

pressure to preserve different segments, a communicational pressure that follows from information-theoretic constraints (Shannon 1948). I further show that the communicational force that predicts which segments will lenite does not follow the pattern that has been observed for merger avoidance (e.g. Wedel, Jackson, & Kaplan 2013, Wedel, Kaplan, & Jackson 2013), but a different pattern, which predicts several duration-related phenomena such as word length (Piantadosi et al. 2011), word duration (Seyfarth 2014), and segment duration (Cohen Priva 2015).

I begin by reviewing the context of the actuation problem and by separating the goal of this article—predicting which language undergoes which lenition—from other aspects of sound change processes. In §3 I introduce the problem by discussing word-final lenition in Indonesian, American English, and Spanish. Section 4 outlines challenges for providing information-theoretic solutions to the actuation of lenition processes and proposes a solution, which is then tested across several corpus studies in §5. Section 6 discusses the merits and limitations of using existing functional-load accounts to solve the same problem.

2. Background.

2.1. Sound change.

Weinreich et al. 1968 divided the understanding of sound change into several related PROBLEMS. One of them, the ACTUATION PROBLEM, deals with the challenges associated with predicting why some languages undergo certain sound changes while others do not, and why changes occur when they do. What makes one language preserve a sound while another language lenites it? The following is a brief review of some existing approaches to the varying concerns within the actuation problem, followed by my proposal.

Several acknowledged principles determine which sound change processes are more likely than others. Some sound changes are more phonologically or phonetically plausible than others. For instance, /p/ is more likely to change into /b/ in a given language (voicing), and less like to change into /n/ (nasalizing, voicing, and changing place of articulation). Similarly, sound systems often maximize contrast between sounds while minimizing effort (Lindblom 1986, Lindblom & Maddieson 1988, Flemming 2004). Thus several /q/-lenition processes in Arabic (Palva 1965, Watson 2002) seem to maintain contrasts by leniting /q/ differently in different dialects (1).

(1) Surface realization of /q/, /ɡ/, /k/, and /ʔ/ in several dialects of Arabic. Within each dialect there is no overlap between the realization of /q/ and the realizations of other phonemes.

DIALECT/PHONEME	/q/	/ɡ/ or /dʒ/	/ʔ/	/k/
Modern Standard Arabic	[q]	[dʒ]	[ʔ]	[k]
Galilean Arabic	[k]	[dʒ]	[ʔ]	[tʃ]
Egyptian Arabic	[ʔ]	[ɡ]	deleted	[k]
Jordanian Arabic	[ɡ]	[dʒ]	[ʔ]/deleted	[k]

/q/ lenites to [k] in Galilean Arabic, which has no /k/; to [ɡ] in Jordanian Arabic, which has no /ɡ/; and to [ʔ] in Egyptian Arabic, in which /ʔ/ is usually deleted. Martinet (1952) used related principles to explain how an individual sound change can lead to chain reactions by causing undesired confusability between existing sounds or by leaving perceptual gaps that can be filled by other sounds. The process of chain shifting has been used to explain both historical and ongoing vowel shifts (Wolfe 1972, Labov 1994, Langstrof 2006, Campbell 2013). These theories can explain the course of sound changes, but not what triggered the initial change in otherwise stable languages.

One commonly assumed trigger is the existing variability in language production. Speech is perceived in a noisy environment, which will necessarily lead to variability in both production and perception (Hockett 1965, Blevins 2006, Kingston 2007). Sound adjacency will lead to overlapping gestures and coarticulation, causing sounds to vary from their ideal production in systematic ways (Hockett 1955, Stevens & House 1963, Öhman 1966, Hillenbrand et al. 2001). Coarticulation and other systematic variations in production that may be misinterpreted by the listener are taken by Ohala (1993a) as preconditions for sound change. Because language is communicative, many researchers have also agreed that listener perception necessarily plays some role in sound changes (Ohala 1993b, Lindblom et al. 1995, Hale 2003, Blevins 2006). Whether in production or perception, variability is accepted as a precondition of language change, and the presence of variability can explain how stable languages can change. However, since variability occurs in all languages, such theories do not currently provide a direct answer to the actuation problem: why variability leads to particular changes in different languages at particular times. Answers to this question have even been labeled 'fruitless pursuits' in Ohala 1993a, and many theories instead focus on the set of possible changes rather than investigate their spread (Hale 2003).

Past work that attributed the spread of sound change to social factors (Labov 1965) has made some headway into this aspect of the actuation problem: Milroy and Milroy (1985) noted that sound change spreads in a regular way throughout communities from group to group, and Labov (2001) proposed that influential people are more likely to initiate sound change with a speech variant. In recent years considerable attention has been given to the role of individuals in the actuation of sound change (see Stevens & Harrington 2014). On the one hand, Baker and colleagues (2011) argue that sound change will be actuated if there is a difference in variability in the pronunciation of certain segments for groups of individuals. For example, in English there are speakers who may be classified as /s/-retractors (more likely to produce /s/ as [ʃ]) and speakers who may be classified as non-/s/-retractors. They hypothesize that sound change occurs when a listener from one group interprets an extreme variation produced by an influential member of the other group as a different speech target and adjusts their own production. On the other hand, Garrett and Johnson (2013) and Yu (2013) argue that it is production instances which are relatively close to the ideal target that drive sound change. Slight deviations from the expected ideal target or ambiguous productions may be accepted into a listener's concept of the target, which can lead to a gradual shift in the category average and eventual sound change. Productions too far from the mean will be rejected. Garrett and Johnson (2013) argue that individuals who attach social significance to coarticulation or variation are more likely to drive sound changes, while Yu (2013) claims that it is individuals with a low autism quotient who fail to accommodate for coarticulation and thus drive the change. By isolating the type of variability and the individuals more likely to spread change, new actuation accounts are able to better identify which sound changes are likely to occur and to better explain how these changes spread.

One of the issues that sound change accounts need to address is that sound change is relatively rare: languages tend to be stable and change slowly. For most of the accounts discussed it is assumed that there is universal variability of sounds, affected by human error and coarticulation, and that the natural ability of humans to compensate for these errors prevents change (e.g. Hockett 1965, Ohala 1993b, Blevins 2006, Kingston 2007, among others). Similarly, in the new actuation proposals of Baker and colleagues (2011), Garrett and Johnson (2013), and Yu (2013), it is the infrequency of many factors that have to cooccur (variable pronunciation, ideal social conditions, update of perception, repetition of a variant) that lead to the relative absence of sound changes.

In contrast, some accounts argue for the existence of preserving forces that do not rely on phonetic factors. Instead, the communicative function of language is responsible for the absence of change. The more a sound change would lead to loss of information, the less likely it is to occur. Hume (2008) argues that high predictability leads to instability, which in turn leads to sound change. Cohen Priva 2012 argues that languages balance information with effort, and segments that provide too little information to justify their effort (information measured as INFORMATIVITY; Cohen Priva 2008, Piantadosi et al. 2011) are more likely to be affected by weakening. Campbell (1996) and Blevins and Wedel (2009) argue that exceptions to regular sound changes can occur if the change causes new homophony in certain cases. Wedel, Kaplan, and Jackson (2013) used a corpus of mergers to show that functional load (measured in number of minimal pairs, but see Hockett 1955, Surendran & Niyogi 2006 for other definitions) correlates with the frequency of merger: segments with high functional load are less likely to merge. Further work by Wedel, Jackson, and Kaplan (2013) showed that the result was more robust for the lemma form of words, for words in the same syntactic category, and for words with similar frequencies. Sóskuthy (2015) modeled hypothetical sound change and showed that using a combination of contrast, phonetic biases, and functional load produces a more accurate model.

Previous work has delved deeply into the question of why and how languages change. Preservation accounts solve a different aspect of the actuation problem: they attempt to predict which sounds will change in a particular language. In the accounts of Garrett and Johnson (2013) and Yu (2013), it is very difficult to predict which sounds will change, but easy to explain how a change was actuated and spread. The account of Baker and colleagues (2011) makes it possible to analyze sound variation in order to potentially predict possible changes, but it does not explain, outside of idiosyncratic phonetic reasons, why that pattern of variation may exist initially. Preservation theories are a means to determine and explain which sounds will become susceptible to variation that could lead to change.

2.2. LENITION. There are multiple, often-conflicting definitions of what constitutes lenition (see e.g. the introduction of Bauer 2008). This article is agnostic about the underlying mechanism of lenition, and it takes the term *lenition* to mean the set of processes typically described as ‘lenition processes’ (summarized in Kirchner 2004: 313). Therefore, the following is not meant to be an exhaustive review of the existing research, but rather to restate the range of processes that are commonly referred to as lenition processes and that the current proposal is meant to address.[2]

Lenition, as the name implies, is the WEAKENING of a consonant in an active phonological process or a historical sound change. Most authors include degemination, debuccalization, spirantization, voicing, approximation, and deletion (e.g. Kirchner 1998, Lavoie 2001, Gurevich 2004, Bauer 2008). Some authors include additional processes under the lenition umbrella term, but these extensions typically depend on what the authors of such papers regard as the formal or functional mechanism that underlies lenition.

Some accounts (e.g. Kirchner 1998) view lenition as an effort-reduction process (but see Kaplan 2010 for a criticism of this approach). Other accounts treat lenition as a product of hypoarticulation (Lindblom 1990, Bauer 2008): when speakers hypoarticulate they may fail to produce stop closure, allow obstruents to passively voice in intervocalic environments, or fail to produce a segment altogether. Hypoarticulation and effort-reduction accounts are mostly compatible.

Both explanations require a triggering component to account for their actuation in a sound change process. Every lenition process applies to segments that are maintained in

[2] See Kirchner 1998:Ch. 2 and Bauer 2008 for a more comprehensive review.

other languages: regardless of whether speakers attempt to reduce effort or hypoarticulate, one needs to explain why such processes occur in one language rather than another. This problem is particularly troubling considering that there are languages that repeatedly lenite some sounds and repeatedly preserve others, without a clear phonetic motivation (see §3).

I consider lenition processes that are mostly uncontested: degemination, debuccalization, spirantization, voicing, approximation, and deletion. The focus in this article is on why specific segments in each language become the targets for repeated lenition. Lenition processes that affect a natural class of segments (e.g. spirantization of all stops, debuccalization of all segments in codas) are not included in this analysis. When a natural class of segments undergoes lenition, the reason may stem from properties of the individual segments, but it is more likely to stem from the properties of the natural class and its defining features. Since the properties of natural classes are different in nature from segmental properties, such processes are outside the scope of this article.

2.3. PROPOSAL. I argue here that highly informative segments are more likely to be preserved and less likely to undergo lenition. In this account, segments will be prone to lenite if they do not provide enough information in a particular language. The proposal is simple. (i) Maintaining segments faithfully is an active process in a language, motivated by the need to transmit information. (ii) The cost languages pay to faithfully produce segments varies by segment and phonological environment, as such cost would depend on phonetic factors (articulatory and perceptual). (iii) Sounds whose information-based contribution is too low to justify the cost paid to maintain them are less likely to benefit from preservation pressures and therefore more likely to lenite.

This account is functionally motivated for both hypoarticulation and effort-reduction accounts. The need for adequate information can be taken as a reason to refrain from hypoarticulation that would lead to lenition. Alternatively, the cost of maintaining faithful transmission can be seen as effortful, and low information would lead to lower effort and less effortful outputs. The account is compatible with all lenition accounts, and I remain agnostic in this article about the mechanism behind lenition. The motivation here is only that low relative amounts of information will lead to less pressure to maintain faithfulness (and therefore greater pressure to lenite).

I consider two related measurements of information: FREQUENCY and INFORMATIVITY. Frequency has been shown to affect the duration and lenition of words and segments (Zipf 1929, Bybee 2000, Bybee & Hopper 2001, Pierrehumbert 2001, Bell et al. 2009), as well as word length (number of segments; Zipf 1935). Informativity—that is, the mean predictability of segments (measured in bits of information)—has also been shown to have an effect on word length, segment duration, and deletion rates (Cohen Priva 2008, 2015, Piantadosi et al. 2011, Seyfarth 2014). Several of these studies found that informativity successfully replaces frequency as an explanatory factor (e.g. Piantadosi et al. 2011, Seyfarth 2014), suggesting that information-based constraints may be the underlying cause for frequency-related phenomena. I define the methods I used for these measurements, following Cohen Priva 2008, Piantadosi et al. 2011, and Seyfarth 2014, in §5.

3. WORD-FINAL LENITION IN INDONESIAN, SPANISH, AND ENGLISH.

3.1. OVERVIEW. In this section I present three cases of lenition of different segments in different languages. These processes exemplify the challenge of solving the actuation problem in the context of lenition. The two main issues are that (i) different languages lenite different segments, and (ii) there are languages that seem prone to lenite certain segments, as these segments repeatedly lenite, either in multiple environments or in multiple varieties of the language. Neither issue is explained straightforwardly in current approaches to sound change. I focus on the problem of /k/-lenition in Indonesian, /t/-lenition in English, and /s/-lenition in Spanish.

3.2. Indonesian /k/-lenition. Word-final and coda stops in Indonesian are often unreleased (Soderberg & Olson 2008). /k/ is affected by a more extreme version of the general lenition pattern and surfaces as [ʔ] in such environments, a process other voiceless stops, and in particular /t/, do not undergo. /s/ is allowed in word-final positions and is not affected by the aforementioned lenition process. Indonesian does force phonotactic constraints even on loanwords, on clusters for example, but such constraints do not forbid word-final /s/: for example, *kelas* 'class'.

The case of Indonesian word-final /k/-lenition is surprising from the phonetic standpoint. Although it may be expected for /k/ to lenite in cases in which /t/ and /s/ do not, Indonesian /k/ also lenites in environments in which /p/ is preserved. Ohala (1983) informs us that /p/ is the least audible of all voiceless stops, which correlates with its greater absence from inventories of voiceless stops than /k/ (Sherman 1975) and its being less frequent than /k/ crosslinguistically (Maddieson 1984, Moran et al. 2014). Indonesian has a /k/-lenition process, but does not have a parallel /p/-lenition process. If languages preferentially lenite their phonetically marked segments, what would make Indonesian lenite /k/ and not /p/?

One solution is that deletion is defined not by phonetic markedness but by phonological markedness. In de Lacy 2002 and related work, dorsals are more marked than labials. Thus, despite /p/ being less frequent crosslinguistically and less audible, by virtue of being a labial it is less marked than /k/ and therefore less likely to undergo word-final lenition. This solution solves the problem of Indonesian /k/ lenition specifically,[3] but does not bear on the two following lenition processes.

3.3. American english /t/-lenition. American English word-final /t/ deletes at varying rates (Guy 1991), intervocalic /t/ is tapped (Kahn 1976, Zue & Laferriere 1979), and word-medial /t/ is more likely to delete than other voiceless stops (Cohen Priva 2015). Similar patterns are observed across many varieties of English. In several varieties of British English, /t/ is the target of several different socially conditioned lenition processes in intervocalic environments. The most famous and widespread pattern is debuccalization (Mathisen 1999), in which /t/ surfaces as a glottal stop. In Irish English varieties /t/ surfaces as [ɾ], [ʔ], [h], [t̞] (an apico-alveolar fricative) or deletes altogether (Hickey 2004). Similarly, in West Midlands English varieties /t/ may surface as [ɾ] or [ʔ] (Clark 2004).

English /k/ and /s/ do not undergo the lenition processes that affect /t/ in the same environments. The absence of /s/-lenition is evident in the number of underlying word-final /s/ sounds that do surface in corpora such as the Buckeye corpus (Pitt et al. 2007): only 1.6% of underlying word-final /s/ sounds did not surface as [s] or some other strident: 94% surfaced as [s].[4] Similarly, only up to 3.5% of all /k/ sounds did not surface as some velar stop, and 92% surface as [k]. In contrast, almost 11% of /t/ do not surface as any coronal or the glottal stop, and only 36% surfaced as [t].

The accumulation of /t/-lenition processes is surprising. Coronals are less marked than dorsals and labials in any account, and preserving /p/ and /k/ while deleting /t/ is

[3] Several authors rank labials as more marked than dorsals; see discussion in de Lacy 2002:§5.3.3.4.

[4] Calculated using the word-level files of the Buckeye corpus, using comparison between the last segment of the underlying representation and the last segment of the surface form.

predicted to be impossible (de Lacy 2002:§5.3.3.2). In surveys of lenition processes (Kirchner 1998, Gurevich 2004), processes that target only coronals are quite rare. Only English targets solely coronals in intervocalic contexts, only one other language deletes solely coronal stops word-finally (Umbrian, but see Buck 1904:146 for word-final /k/ deletion), and only two other languages (Middle Egyptian, Limbu) have some other word-final process that targets solely coronals. Even if the hundreds of languages included in the survey are not fully representative of all the world's languages, it is unlikely that they are biased toward omitting /t/-lenition processes. If the range of lenition processes that target segments in a particular language is sampled from the distribution of all possible sound change processes, the coincidental reemergence of typologically rare processes in any one language is unexpected. Rather, it is expected that if some variety of English lenites its /t/, other varieties would weaken their more marked segments, such as /p/ or /k/.

It should be noted that contact between varieties can account for some spreading of lenition processes, but not for all. Varieties that lenite /t/ to [ɾ] preserve the place of articulation but not the manner, while varieties that lenite /t/ to [ʔ] preserve the manner but not place of articulation. Both must be traced to varieties in which /t/ surfaces as a coronal stop, a [t]. If a variety copies a lenition process, why copy the sound undergoing lenition but not the outcome of the lenition? Similarly, the existence of /t/-lenition in intervocalic contexts does not predict /t/-lenition in codas. Both the triggering environments and the outcomes are different in such cases. Many of the processes are therefore independent from one another, or stem from a cause yet to be discovered.

3.4. SPANISH /s/-LENITION. Several Western Romance languages undergo or have undergone /s/-lenition processes word-finally and in syllable codas, usually with an intermediate /s/ → [h] stage, followed by deletion (Harris 1969, Gess 2001, MacKenzie 2010, Sauzet 2012). I focus on Spanish, as many of its dialects exhibit some level of /s/-deletion synchronically (Terrell 1979, Poplack 1980, Hochberg 1986, Morris 2000, Carvalho 2006). In the LDC corpus of /s/-deletion (Fox 2001), which was derived from the Spanish CALLHOME corpus (Canavan & Zipperlen 1996), the least-deleting dialects delete 10% of their /s/s (Spain, Colombia), while other dialects delete at least 20%, with Puerto Rican Spanish /s/-deletion, studied in Poplack 1980 and Hochberg 1986, at 67%.

Like English /t/-lenition, /s/-lenition is surprising from the markedness point of view, as /s/ is extremely frequent crosslinguistically (73% of the languages in Maddieson 1984 have some voiceless coronal sibilant) and a coronal. Spanish does not have word-final /t/ or /k/ in its native vocabulary, but word-final /t/ and /k/ are not forbidden by Spanish phonology. Thus *robot* and *Internet* are pronounced with final /t/ sounds, and *zinc* (*cinc*) and *New York* (*Nueva York*) are pronounced with final /k/ sounds. Other patterns, such as particular consonant clusters that are not allowed in Spanish, are eliminated through epenthesis and deletion: for example, *standard* is borrowed as *estándar* (the original meaning of 'banner' is *estandarte*).

3.5. THE ACTUATION OF LENITION PROCESSES. Indonesian, English, and Spanish each lenite a segment that is not lenited by the other two languages. Two of the lenition patterns are unusual and yet repeat across several of the environments and varieties of the language (English) or across several related varieties (Spanish) and languages (Western Romance). What predicts which language will lenite which segment? It is not immediately clear which phonetic or phonological factors can predict the language-specific patterns. In the following sections I propose using information preservation (and lack thereof) to predict the observed lenition patterns.

4. Information-theoretic solutions.

4.1. Overview of information-theoretic research in linguistics. There is a growing body of research on the role of information-theoretic constraints in human language. As predicted by information theory (Shannon 1948), speakers do not provide too little or too much information at a given time (Aylett & Turk 2004, Levy & Jaeger 2007, Jaeger 2010; see Jaeger & Buz 2017 for a comprehensive review). The preservation-of-information rate is obtained by omitting, reducing, or hypoarticulating low-information linguistic material or by expanding or hyperarticulating high-information linguistic material (Lindblom 1990). Optimization can also occur at the level of an entire language. Word length (number of segments) is highly correlated with how predictable words are on average (word informativity; Piantadosi et al. 2011), more than with how frequent they are (though frequency is highly correlated with average predictability; see discussion in §5.2). Similar principles can be used to predict which language will be affected by which sound change process.

Current research on information-theoretic factors affecting linguistic production predicts that uninformative forms would be reduced (shorter, less distinctive, or elided) relative to equivalent informative forms. Such accounts can therefore predict lenition if the segments that undergo lenition provide less information in the languages in which they lenite than in languages in which they do not. To make this prediction concrete it is necessary to tackle several issues. First, it is necessary to define what it means to provide information at the segmental level. Second, the information-theoretic approach must explain why lenition processes are often 'exceptionless' (in the Neogrammarian sense; Hale 2003), in that once a lenition process comes to exist in a language, it may apply indiscriminately, not just in contexts in which that segment provides little information. Finally, the approach must be flexible enough to be capable of predicting the lenition not only of the least informative segments, but also of informative ones. The rest of this section addresses these three issues.

4.2. Local predictability. Information-theoretic accounts define information as surprisal: the less predictable the message, the more information it provides. To assess the amount of information provided by some event in some context, the negative log probability of observing the event given the context is taken, as in 2.

(2) Surprisal: $-\log\Pr(\text{event}|\text{context})$

When the context is missing, the result is 3, an information-theoretic form of the frequency of the event, connecting frequentist accounts and probabilistic accounts.

(3) Frequency: surprisal without context: $-\log\Pr(\text{event}|\emptyset)$

Frequency-based accounts do not necessarily rely on information-based reasoning. Zipf (1929) argues that frequent segments are articulated more than other segments and are simplified by usage. Bybee (2000) and Pierrehumbert (2001) use frequency directly in their exemplar models. Both approaches can be explained in closely related information-theoretic terms: frequent events (words, segments) are less informative, everything else being equal, and are therefore under a greater pressure to be reduced, in duration or articulatorily.[5]

In linguistic research, the event in 2 is usually the appearance of some linguistic material (e.g. segment, word), and the context is what is already known in the utterance, as defined by some model of communication. In the model I use (following van Son &

[5] Other frequency effects such as 'entrenchment' (Pierrehumbert 2001) cannot be modeled in information-theoretic terms, but these are not the focus of this article.

Pols 2003), the information provided by the /k/ in the word *lack* is computed using all of the previous segments within the word as context (4). Other models may define context differently and can be empirically compared to the van Son and Pols (2003) definition of context used here.

(4) $-\log \Pr(/k/|/læ\text{-}/)$

Using conditional probability allows for the comparison of the information provided by a segment in different contexts. The /k/ in *lack* provides a lot of information if calculated as in 4, since other sounds such as /s/ are much more likely to follow in that context. For instance, the /s/ in *last* follows /læ-/ much more frequently than /k/ does (*lack* appears fifty-nine times in the Switchboard corpus; *last* appears 2,172 times). Thus the probability of observing /k/ in this context is small, and the amount of information it provides when it appears in that context is high. In contrast, /k/ is the only sound that can follow /kənɛtɪ-/ (e.g. in the word *kinetic*). The probability of /k/ following /kənɛtɪ-/ is therefore 1, and the negative log of 1 is zero: /k/ provides no information in the context of /kənɛtɪ-/; it is completely redundant.

In order to calculate how predictable a segment is in some context (its local predictability), it is possible to use counts (as implied above for *lack* and *last*).[6] The local predictability of a segment in context is estimated as the number of times the segment appeared in that context, divided by the number of times that context appeared with any segment (5).

(5) Segmental local predictability: $\Pr(\text{segment}|\text{context}) := \dfrac{\#(\text{segment in context})}{\#(\text{context})}$

For example, if we assume that in our corpus only the words *lack* and *last* begin with /læ-/, and that /k/ follows /læ-/ fifty-nine times and /s/ follows /læ-/ 2,172 times, the local predictability of observing /k/ after /læ-/ is 59 (the number of times /k/ followed /læ-/) divided by 59 + 2,172 (the number of times /læ-/ appeared in the corpus), as in 6. This yields 0.026. Taking the negative log of this number using log base 2 yields 3.63 bits of information, as in 7.[7]

(6) $-\log_2 \dfrac{\#(\text{occurrences of /læk/})}{\#(\text{occurrences of /læ–/})}$

(7) $-\log_2 \dfrac{59}{59+2172} = 3.63$

Local predictability at the segmental level is well studied. Van Son and Pols (2003) and van Son and van Santen (2005) found that the redundancy of a segment, measured using two different but related methods, was a significant factor in predicting the duration of segments, even after taking prosodic factors into account. Van Son and Pols (2003) measured segment redundancy using the preceding context within a word, as defined in this article. Van Son and van Santen (2005) used the negative log probabilities of classes of segments appearing in strong and weak positions for redundancy estimates. Local predictability is therefore the initial answer to the question of how to measure information at the segmental level: following information theory and previous research, it is possible to measure how predictable a segment is in each context in which it appears, thus predicting how reduced it is likely to be in this context. Two issues remain: predicting 'exceptionless' properties of sound change, and predicting the lenition of highly informative segments.

[6] This is a maximum likelihood approach.

[7] Base 2 is standardly used in information theory.

4.3. Informativity. Local predictability changes from one context to another and therefore cannot predict exceptionless sound change. It predicts that sounds will lenite only in contexts in which they are predictable, but Cohen Priva 2008 demonstrates that this is not the case: /p/ and /k/ tend to be preserved in English even when uninformative, while /t/ tends to be reduced even when informative. Cohen Priva 2008 argues that the solution for word-medial deletion of unpredictable /t/ and lack of deletion for redundant /p/ and /k/ lies in informativity, or average predictability. Thus, segments that are usually predictable will have low informativity even in contexts in which they are unpredictable, and segments that are usually unpredictable will have high informativity even in cases in which they are predictable or even redundant. Such principles were subsequently demonstrated to hold even at higher linguistic levels: word length (measured in number of sounds) seems to correlate with informativity crosslinguistically, more than with frequency (Piantadosi et al. 2011). Words and segments that have low informativity have shorter duration even when word length, phonetic properties, frequency, and local predictability are controlled for (Seyfarth 2014, Cohen Priva 2015).

The informativity of a segment is calculated by averaging the number of bits that segment provides in each and every context in which it appears (again, in this model context is preceding segments in the same word, following van Son & Pols 2003). Averaging across contexts takes into account how frequent that context is, as in 8. Calculating informativity using this method therefore necessitates a phonemic lexicon of the language and word counts.

(8) $$\sum_{contexts} \frac{\text{Pr(segment appears with context)(information provided by segment in context)}}{\text{Pr(segment)}}$$

Combining 8 with 5, the information measurement for each context, yields 9, which can be simplified to 10, the informativity of that segment in the language (the expected value of its self-information).

(9) $$\text{Informativity: } -\sum_{contexts} \frac{\text{Pr(segment appears with context)logPr(segment|context)}}{\text{Pr(segment)}}$$

(10) $$\text{Informativity (simplified): } -\sum_{contexts}\text{Pr(context|segment)logPr(segment|context)}$$

4.4. Predicting lenition for informative segments. Most studies of information-theoretic effects in duration or reduction predict variable reduction rates but do not attempt to predict the baseline duration and reduction. For instance, Jurafsky and colleagues (1998) compare the duration of frequent words in predictable and unpredictable contexts but do not attempt to predict the baseline duration of the individual words: predictable *that* is compared to unpredictable *that*, not to other function words. Levy and Jaeger (2007) compare the insertion and omission of *that* in predictable and unpredictable contexts, but do not attempt to predict the base insertion/omission rate, or why *that* can be omitted more than other function words. It is difficult to extend this approach to the prediction of lenition: for example, even if /t/ usually has lower informativity than /f/ in languages that have both sounds, it should not necessarily be expected to lenite in every language. Perhaps phonetic reasons allow /t/ to have lower baseline informativity than /f/. Such differences in the least amount of information that would lead to the preservation of a segment predict that some sounds may begin to lenite even when they are still quite informative: /t/ may begin to lenite if it provides less than two bits of information, but /f/ may lenite if it provides less than three. Therefore, the proposed account can predict that sounds that are usually highly informative may lenite when their informativity is relatively low, not just when it reaches some absolute low value. Cohen Priva 2012 argues that informativity justifies articulatory and perceptual effort, but without a plausible way to measure effort, it would be difficult to predict which sounds will lenite. As I stress in §2.3, and unlike in Cohen Priva 2012, I do not argue that the phonetic baseline necessarily stands for articulatory effort. Rather, the phonetic baseline stands for the cost languages pay to guarantee faithful transmission of a segment. It may be interpreted as articulatory effort, but also as being at risk for misperception, or being highly affected by hypoarticulation.

I propose two alternative approaches to controlling for the phonetic baseline of each sound. One approach is to use the information similar sounds provide in other languages as a baseline in order to assess when a language deviates from that baseline, as in §5.1. The other approach is to control for phonetic factors directly in a multiple regression (following Cohen Priva 2008, Bell et al. 2009), as in §5.3.

5. USING INFORMATIVITY TO PREDICT WORD-FINAL LENITION.

5.1. USING INFORMATIVITY TO PREDICT LENITION. Cohen Priva 2015 showed that if phonetic factors are controlled for, word-medial American English consonants are shorter and more likely to delete when their informativity is relatively low. Extending this idea to word-final lenition predicts that for a given segment, duration reduction, as well as deletion, will increase if its informativity is low. This leads to two methodological alternatives, both of which are explored in the following sections. The first approach is to use crosslinguistic comparison to estimate how many bits of information a segment SHOULD provide. If a segment provides less than the expected amount of information given the information of the segment crosslinguistically (taking the information of other segments in the language into account), it will be predicted to be under pressure to lenite. This follows from a straightforward extension of Cohen Priva 2015: if some segment provides two bits of information in one language and four bits in another language, it will be under greater pressure to lenite in the two-bit language than in the four-bit language.[8] By extending this approach to more than two languages it is possible to estimate if a segment's informativity is unusually low in some language. The second approach is to control for phonetic factors in a multiple regression and see whether informativity predicts the preservation of segments word-finally. The second approach is applicable only if segment-level annotations are available.

In the following sections I use both methods to predict word-final lenition. Since accurately annotated corpora are not available for Indonesian and Spanish, I first explore the crosslinguistic approach to see whether the proposed method predicts the lenition of Indonesian /k/, American English /t/, and Spanish /s/. I do so by comparing the informativity of /k/, /t/, and /s/ crosslinguistically. I then use the same methods and corpus used in Cohen Priva 2015 in order to predict word-final deletion in American English.

5.2. CROSSLINGUISTIC COMPARISON.

MATERIALS FOR CALCULATING INFORMATIVITY. I calculated the informativity of segments in several languages in order to estimate whether the informativity of /k/ in Indonesian, /t/ in American English, and /s/ in Spanish is unusually low. Informativity was calculated as explained above, following the procedure detailed in Cohen Priva 2015. If a corpus included both spoken and written uses, only spoken frequency counts were used. The reported word types are only of words that occurred at least once in the corpus. If the lexical entry contained duplicate transcriptions (e.g. two different pronunciation alternatives), the first entry was used.

In all of these studies I rely on phonemes rather than on surface phonetic representations, for several reasons. First, surface forms are varied even in the same environment (see also discussion in Cohen Priva 2015:253–54 and Appendix A). Second, lenition

[8] Several necessary adjustments to this strong prediction are beyond the scope of this article. Those can include the consideration of language-specific phonetic factors such as perceptual competition, number of segments, and syllable structure.

processes include deletion: it is not clear if information should be assigned to missing elements, and if so, whether all missing elements should be treated as identical regardless of underlying representation. Finally, the goal of this article is to understand the reasons leading to the actuation of a phonological process, for example, /s/ to [h]; assigning properties to [h] rather than /s/ presupposes the process that needs to be explained.

One may wonder whether the informativity and frequency can be reliably estimated from the given corpora. No corpus is representative of the entire language due to differences in subject matter, genre, register, and more. By using mostly spontaneous speech corpora we hope to approximate language as used by speakers in regular conditions. A different concern is raised by Baayen (2001) and Daland (2013), who point to properties of language that make small data sources unreliable samples even for the genre and register to which they belong. I addressed these concerns by checking whether informativity values rely too heavily on low-frequency words and items. This was done by comparing the correlation between informativity values in the entire corpus with informativity values in a minimally different corpus from which words that appeared fewer than four times were excluded. For all languages used in the corpus, the Pearson correlation coefficient was extremely high (> 0.98). The consistency of informativity is not surprising, as it is an expected value over the entire lexicon and is therefore less susceptible to phenomena originating from reliance on smaller data sources.

(i) Indonesian: All articles from the Berita Satu Indonesian newspaper[9] were downloaded and converted to text. Punctuation was stripped from the text and only lowercase words were included, resulting in around 3 million words. The words were used to calculate word counts and infer phonemic representation. The Indonesian alphabet is mostly phonemic, making the task relatively straightforward, with digraphs for /ɲ/, /ŋ/, /ʃ/, and /x/ changed to their respective underlying segments. This is the only written corpus used in this set of studies. The number of word types was 36,807, the number of word tokens was 3,186,933, and the number of segment tokens was 18,820,822.

(ii) American English: Word counts were taken from the Switchboard (~400 speakers in ~2,000 conversations, on ~70 general topics; Godfrey & Holliman 1997), Fisher (~12,000 participants in ~12,000 conversations, on 100 topics; Cieri et al. 2004, Cieri et al. 2005), and Buckeye (forty interviews with different speakers; Pitt et al. 2007) corpora. The *CMU pronouncing dictionary* (Weide 2008) was used for words' underlying representations. The number of word types was 42,121, the number of word tokens was 23,513,056, and the number of segment tokens was 71,700,964.

(iii) Spanish: Word counts were calculated using the training section of the LDC CALLHOME Spanish transcripts (Wheatley 1996), and lexical phonemic information was taken from the LDC CALLHOME Spanish lexicon (Garrett et al. 1996). Word counts include eighty unscripted conversations between native speakers. The corpus is mostly phonemic, and allophonic variations in manner for /b/, /d/, /ɡ/, and /s/ (spirantization for the stops, voicing for /s/) were collapsed to a single underlying representation. The number of word types was 9,090, the number of word tokens was 143,086, and the number of segment tokens was 530,405.

For other languages, I used every LDC CALLHOME and CALLFRIEND lexicon that provided phonemic data. These corpora consist of unscripted and uninstructed conversations between two speakers of the relevant languages. These include:

[9] http://www.beritasatu.com/, downloaded May 25, 2016.

(iv) ARABIC: The LDC Egyptian Colloquial Arabic lexicon (Kilany et al. 1997) was used for counts and phonemic data. Word counts were derived from 100 conversations. Geminates were treated as a single segment, followed by a special lengthening marker. The number of word types was 16,272, the number of word tokens was 153,330, and the number of segment tokens was 886,241.

(v) JAPANESE: The CALLHOME Japanese lexicon (Kobayashi et al. 1996) was used for counts and phonemic data. Word counts were derived from eighty conversations. Geminates were treated as distinct from singletons (rather than lengthened singletons), following Kawahara 2016 and advice from Natasha Warner (p.c.). The number of word types was 9,265, the number of word tokens was 748,482, and the number of segment tokens was 1,715,874.

(vi) MANDARIN CHINESE: The CALLHOME Mandarin Chinese lexicon (Huang et al. 1996) was used for counts and phonemic data. Word counts were derived from eighty conversations. The number of word types was 6,344, the number of word tokens was 155,192, and the number of segment tokens was 764,191.

(vii) KOREAN: The Korean telephone conversations lexicon (Han et al. 2003) was used for counts and phonemic data. Word counts were derived from 100 conversations. Phonemic representation was deduced from the Yale Romanization, with digraphs converted to single segments. The number of word types was 25,213, the number of word tokens was 187,991, and the number of segment tokens was 1,130,277.

I compare here the informativity values of voiceless obstruents, as this is the smallest natural class that includes /k/, /t/, and /s/. More inclusive sets that include all obstruents or all consonants do not lead to different results.

It should be noted that despite having common lexical sources, the informativity profiles of American English and Spanish are quite distinct, as Figure 1 shows, and similarly Mandarin Chinese, Korean, and Japanese have quite different informativity profiles.

INDONESIAN /k/-LENITION RESULTS. The informativity account predicts that if Indonesian weakens /k/ but not similar phonemes, then the informativity of /k/ should be unusually low in Indonesian. This is the case both relative to other segments in Indonesian and relative to the information /k/ provides in other languages. First, Indonesian /k/ provides only 1.81 bits of information (the crosslinguistic average is 2.51), compared to the average for all voiceless obstruents in Indonesian: 3.69 bits; see Figure 2. In addition, no other /k/ in the seven languages sampled has such low informativity, though the /k/ of Korean comes close with 1.83 bits (Fig. 1).[10] The informativity account for lenition therefore predicts that Indonesian would be more likely to undergo /k/-lenition than other languages, as has been observed.

AMERICAN ENGLISH /t/-LENITION RESULTS. American English /t/ provides only 1.35 bits of information (the crosslinguistic average is 2.32), compared to the mean for all voiceless obstruents in English: 3.4 bits; see Figure 3. No other /t/ in the seven languages sampled has such low informativity, with the next least informative /t/ being that of Indonesian, with 1.97 bits (Fig. 1). The informativity therefore predicts the observed /t/-lenition patterns in American English.

[10] See Cohen Priva 2012:Ch. 5 for a discussion of the relatively high frequency of Korean /k/, which could be the reason for its relatively low informativity.

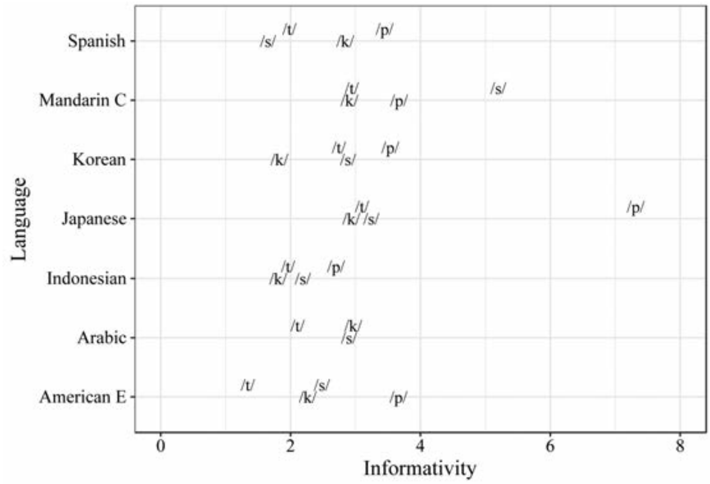

FIGURE 1. Crosslinguistic informativity of /t/, /k/, /p/, and /s/, if available. The y-axis shows different languages, and the x-axis shows informativity, measured in bits of information. Different placements on the y-axis within a language are meant to improve readability.

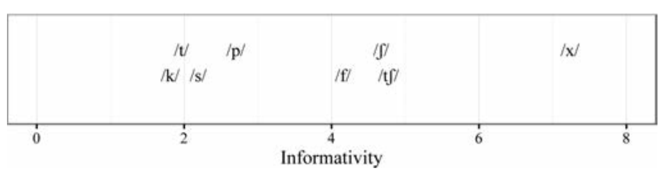

FIGURE 2. Informativity of Indonesian voiceless obstruents, including the semi-peripheral /x/ and peripheral /ʃ/. Informativity is measured in bits of information. Different placements on the y-axis are meant to improve readability.

SPANISH /s/-LENITION RESULTS. Spanish /s/ provides only 1.66 bits of information (the crosslinguistic average is 2.94), compared to the average for all voiceless obstruents in Spanish: 3.37 bits; see Figure 4. As with American English /t/ and Indonesian /k/, no other /s/ in the seven languages sampled has such low informativity, with the next least informative /s/ being that of Indonesian, with 2.2 bits (Fig. 1). The informativity account therefore predicts the observed /s/-lenition patterns.

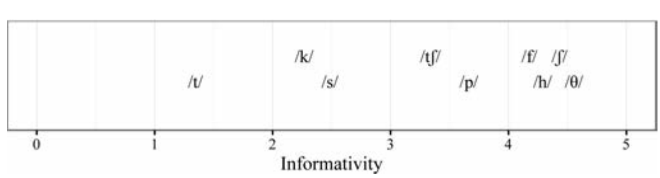

FIGURE 3. Informativity of American English voiceless obstruents. Informativity is measured in bits of information. Different placements on the y-axis are meant to improve readability.

DISCUSSION. In all three languages, the informativity of the segment that is more likely to lenite word-finally is unusually low compared to the informativity of similar segments within the language, and is the lowest informativity for that segment across all of the languages compared in the sample. It is quite unlikely that the match between lenition and low informativity is due to chance. As mentioned above, it is important to stress two caveats. First, having relatively low informativity does not equate to obligatory lenition,

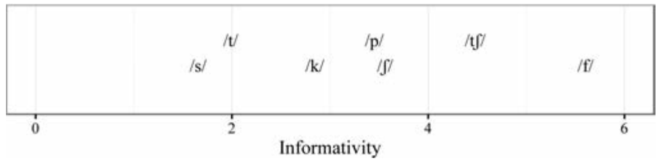

FIGURE 4. Informativity of Spanish voiceless obstruents, including the peripheral /ʃ/. Informativity is measured in bits of information. Different placements on the y-axis are meant to improve readability.

but rather indicates that the segment is under a stronger pressure to undergo lenition. Second, all values must be evaluated relative to similar segments in other languages, since phonetic factors are not controlled for (as opposed to studies in e.g. Wedel, Kaplan, & Jackson 2013, Cohen Priva 2015, and the following study in §5.3).

Initially, it would seem that the results support two related hypotheses. First, it is possible that lenition is correlated with absolute low informativity: all three segments have unusually low informativity values not only compared to their counterparts in other languages, but also compared to all voiceless obstruents in the sample. It seems that most low-informativity obstruents in the sample preferentially lenite.[11] Alternatively, it is possible that lenition is likely to affect segments whose informativity is unusually low relative to that segment's crosslinguistic informativity. While for Indonesian, American English, and Spanish both alternatives predict the observed pattern, the latter alternative seems to account better for other persistent patterns in the data. In all six languages that have /p/, the informativity of /p/ is higher than the informativity of /t/ and /k/. Similarly, in six out of seven languages, /s/ has higher informativity than /t/. Both trends correlate with the distributions of /p/ relative to /t/ and /k/, and of /s/ relative to /t/ (Maddieson 1984, Moran et al. 2014). Thus it seems that even though the results could have argued for lenition on the basis of absolute low values, informativity rankings seem to follow their crosslinguistic distribution, which arguably correlates with their phonetic markedness. If so, phonetically marked segments such as /p/ would lenite well before their informativity falls as low as the segments that are described above (see analysis of /q/-weakening in Cohen Priva 2012:Ch. 3).

In the studies above I did not contrast frequency and informativity directly. Frequency is highly correlated with informativity (in the absence of context, informativity is the same as frequency) and would therefore yield very similar patterns for the languages mentioned above. In the absence of sufficient data to estimate informativity (e.g. in a language for which word-usage counts are not available), frequency can be used as a proxy for informativity.[12] Frequency-based reasoning can follow information-based reasons, but may also follow from other factors, as discussed in Zipf 1929 and Bybee 2000. The following study used multiple regression to predict lenition and is able to contrast frequency and informativity directly.

5.3. VARIABLE WORD-FINAL DELETION IN CASUAL AMERICAN ENGLISH.

INTRODUCTION AND MOTIVATION. The previous studies indicate that informativity is linked with word-final lenition. To further test the relationship between informativity and lenition, it would be useful to control for a range of additional factors that are

[11] Marginal segments are also likely to have very low informativity, as they appear in few contexts and those contexts can be highly predictable.

[12] Frequency and informativity can differ substantially: American English /ŋ/ is rather infrequent (which would predict high informativity), but it appears almost exclusively in very predictable contexts (in *-ing*) and therefore has low informativity.

known to influence lenition. For instance, word frequency is known to affect word- and syllable-level reduction (e.g. Aylett & Turk 2004, Bell et al. 2009). For American English, the lenition pattern in question is word-final deletion. To rule out the possibility that other phonetic factors determine which word-final obstruents are deleted in American English, I conducted a regression-based study of word-final obstruents in postvocalic, preconsonantal (in the following word) environment in American English. The study focuses not just on /t/ but on all obstruents, and it aims to see whether low informativity contributes to deletion after various phonetic and information-theoretic factors have been controlled for.

In essence, this study is a replication of the postvocalic obstruent-deletion study in Cohen Priva 2015, focusing on word-final contexts. The original study investigated word-medial processes, in which /t/-deletion is not grammatically licensed. The following replication therefore extends the previous study to a domain in which phonological /t/-deletion processes exist: speakers do not omit word-medial /t/ in carefully articulated speech, but may omit /t/ word-finally in words such as *just*, even in carefully articulated speech. The choice of postvocalic environments over postconsonantal environments is due to the overrepresentation of /t/ in postconsonantal positions, compared with other obstruents. For instance, in the CMU dictionary (Weide 2008), word-final /t/ in American English appears in word-final consonant clusters 51.2% of the time (type frequency), while word-final /k/ appears in postconsonantal positions only 13.6% of the time.

Methods and materials. I followed Cohen Priva 2015 in using the Buckeye Corpus of Conversational Speech (Pitt et al. 2007), which provides data collected from forty speakers at the Ohio State University conversing freely with an interviewer. The corpus provides several values for each word, including its duration, part of speech, underlying form, and actual pronunciation. For each word, underlying and surface segments were aligned by replicating the procedure detailed in Cohen Priva 2015, such that in cases such as 11, the algorithm would align /b/ with [b], /æ/ with [ɜ], and /s/ with [z], and would regard /k/ as deleted. I restrict the analysis to underlyingly word-final postvocalic, preconsonantal obstruents (the following word begins with a consonant). If a segment did not have a corresponding surface segment, it was considered deleted.

(11) /bæks/ → [bɜz]

As with the previous study, I used word counts from the Buckeye (Pitt et al. 2007), Switchboard (Godfrey & Holliman 1997), and Fisher (Cieri et al. 2004, Cieri et al. 2005) corpora combined to provide overall word counts.

I used several phonological variables to control for base properties of segments (Table 1). In addition to segment-level properties, controls were added for word-level effects as well as the phonological properties of the following segments (Table 2).[13] Following Cohen Priva 2015, and in order to avoid using words with only a few data points, words whose frequency in the data set was less than four were excluded. This resulted in ~13,000 segment tokens from ~450 word types.

I used the step() function (Hastie & Pregibon 1992, Venables & Ripley 2002) in R (R Core Team 2014) to allow the best non-information-theoretic model to be chosen automatically, and then added four information-theoretic variables: word and segment prob-

[13] Some of the features do not correspond directly to phonological features in order to avoid overfitting. English has only one palatal consonant, /j/, which was assigned a dorsal place of articulation and postalveolar binary designation to avoid assigning it features that would apply only to that one segment. Similarly, voicing applies only to obstruents, which contrast in voicing, but not to sonorants, which do not contrast in voicing.

FEATURE	SEGMENTS
Place: Labial	Labials, labiodentals
Place: Coronal	Dentals, alveolars, postalveolars
Place: Dorsal	/k/, /g/ /ŋ/, /j/
Dental	Following segment was /θ/ or /ð/
Postalveolar	/ʃ/, /ʒ/ /tʃ/, /dʒ/, /j/
Stops	/p/, /t/, /k/, /b/, /d/, /g/
Affricate	/tʃ/, /dʒ/
Obstruent	All obstruents
Voiced	Voiced obstruents

TABLE 1. Phonological control variables for base properties of segments.

FEATURE	MEANING
Rate of speech	Measured in lexical phonemes per second
Stress	Preceding vowel has primary or secondary stress
Approximant	Liquids, glides
Lateral	Following segment was /l/
Nasal	Following segment was /n/, /m/, or /ŋ/
Liquid	Following segment was /l/ or /ɹ/
Identical place	Following segment had the same place of articulation

TABLE 2. Additional controls for word-level effects and the phonological properties of the following segments.

ability (its frequency, negative log transformed; see 3), segment informativity (10), and the local conditional predictability of the segment (5). Segment local predictability was residualized using both segment probability and informativity. Thus, it would only be significant if it improved the model beyond the (unconstrained) effect that the variables it is residualized over have.

The model was then reevaluated using a mixed-effects model with the identity of the word and speaker used as random intercepts, as well as with by-speaker random slopes for segment probability, informativity, and local predictability. For a full explanation of the way logistic regressions fit data see Bresnan & Nikitina 2009. The model being trained here follows and uses the exact same packages used in Cohen Priva 2015: lmerTest, lme4, and optimx (Kuznetsova et al. 2014, Bates et al. 2014, Nash & Varadhan 2011, respectively). Additional transformations were taken in order to allow the mixed-effects models to converge. All variables were normalized, which allows for the comparison of effect sizes as well as significance. Finally, multinomial variables were binarized—place of articulation was represented as two binary variables (labial and dorsal), returning true or false, rather than a single three-level variable. I report the coefficients and *p*-values of the variables of interest.

RESULTS. As predicted, high segment informativity predicted lower likelihood to be deleted ($\beta = -0.54$, $SE = 0.17$, $z = -3.219$, $p < 0.01$). Among other information-theoretic variables, low word probability likewise promoted word-final segment preservation, though the effect was marginal ($\beta = -0.2$, $SE = 0.11$, $z = -1.918$, $p < 0.1$), possibly due to the inclusion of word as a random intercept. Segment probability and predictability did not have a significant effect ($\beta = -0.086$, $SE = 0.2$, $z = -0.42$, $p = 0.672$; $\beta = 0.0041$, $SE = 0.08$, $z = 0.051$, $p = 0.959$; respectively).

Among the controls, several variables significantly affected a segment's deletion likelihood. Fast speech rate was correlated with higher deletion rates ($\beta = 0.36$, $SE = 0.052$, $z = 7.044$, $p < 10^{-11}$), and segments that followed a stressed vowel were less likely to be deleted ($\beta = -0.27$, $SE = 0.071$, $z = -3.755$, $p < 0.001$). Stops were more likely to be deleted ($\beta = 0.82$, $SE = 0.13$, $z = 6.415$, $p < 10^{-9}$), as were labials ($\beta = 0.43$, $SE = 0.11$,

$z = 4.014$, $p < 10^{-4}$). Dorsals were less likely to be deleted ($\beta = -0.28$, $SE = 0.14$, $z = -2.089$, $p < 0.05$). No other phonological properties of the segment affected its likelihood to be deleted.

In addition, several properties of the following segment affected the segment's likelihood to delete. Following dentals, postalveolars, and liquids promoted preservation ($\beta = -0.11$, $SE = 0.052$, $z = -2.164$, $p < 0.05$; $\beta = -0.25$, $SE = 0.064$, $z = -3.89$, $p < 0.001$; $\beta = -0.12$, $SE = 0.052$, $z = -2.361$, $p < 0.05$; respectively), while a following approximant promoted deletion ($\beta = 0.21$, $SE = 0.055$, $z = 3.889$, $p < 0.001$).[14] No other phonological property had a significant effect on the likelihood to delete.

Figure 5 visualizes the model estimates for the coefficients of information-theoretic variables and compares them with rate of speech, one of the best predictors of deletion.

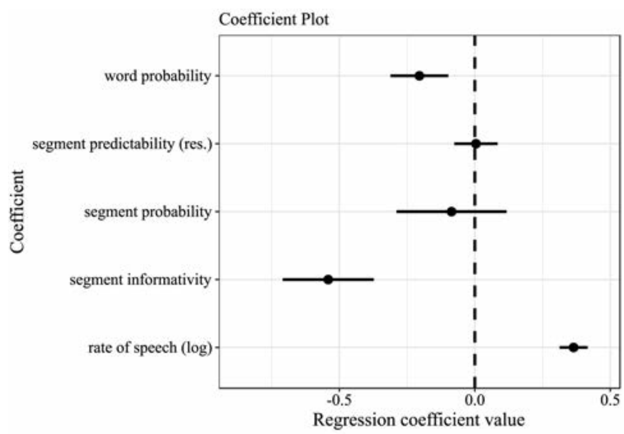

FIGURE 5. The information-theoretic model variables, as well as rate of speech. The plot visualizes the relative estimates and confidence of the different variables. All variables are normalized, which means that the estimates are of the same scale.

DISCUSSION. The results show that informativity is negatively correlated with deletion of word-final consonants in American English, stronger than many phonological predictors such as voicing or dorsal place of articulation. As such, they lend further support to the assumption that segments with high informativity are more likely to be preserved than those with low informativity when a range of phonetic, phonological, and word-level variables are controlled for. The results also show that informativity affects likelihood to delete even in environments in which optional phonological deletion processes exist (e.g. word-final /t/-deletion), and not just in environments in which no phonological deletion processes are licensed, as Cohen Priva 2015 showed.

The information-theoretic variables are particularly intriguing. A segment's probability has no residual effect on its likelihood to delete: it is not the case that frequent segments are necessarily under more pressure to be lenited than infrequent ones (unlike the prediction of Zipf 1929). The lack of effect is due to the inclusion of informativity: informativity explains all of the variance that the segment probability could explain, making segment probability redundant.

The functional motivation for the effect informativity has on lenition is that locally predictable segments are more likely to be reduced. However, only the expected value (informativity) contributed to the prediction of lenition, not local predictability. Thus, the

[14] Since liquids are also approximants, their promotion of preservation should be interpreted as promoting preservation RELATIVE TO OTHER APPROXIMANTS, that is, relative to glides.

results support the Neogrammarian view that sound change is exceptionless: at the stage in which sound change occurs it no longer matters whether a segment undergoing lenition is unpredictable or predictable in context; having relatively low informativity suffices for lenition. The results also support the assumption that informativity, rather than frequency, accounts for lenition. The regression-based study is the first study to contrast frequency and informativity at the segmental level. As previous studies found at the word level (Piantadosi et al. 2011, Seyfarth 2014), once informativity is controlled for, frequency does not contribute to predicting the observed pattern, in this case lenition.

The method used in this section relies on the availability of richly annotated corpora and can be applied to additional languages when similar corpora become available.

5.4. STUDIES SUMMARY. This section explored two methods for predicting the preference to delete some segments rather than others in a language. The first set of studies showed that lenition is more likely to be found for segments whose informativity is unusually low in comparison to the informativity of comparable segments in other languages. The second study showed the effect of informativity on lenition (in this case, deletion) by controlling for the (unconstrained) effect of phonetic and phonological factors directly. Thus, the model determined a baseline for deletion and found that informativity was inversely correlated with deletion, such that low-informativity segments were more likely to be deleted and high-informativity segments were more likely to be preserved. Together, all of the studies above provide converging evidence that implicate informativity in the actuation of selective lenition processes word-finally.

The correlation between low informativity and lenition does not by itself imply causation, and three possible causal accounts are possible. First, low informativity leads to lenition (directly or indirectly). Second, lenition leads to low informativity (again, directly or indirectly). Finally, it is possible that unknown factors are responsible for low informativity as well as for lenition. Of all three, only the first is grounded in previous empirical and theoretical findings. Relationships between low information and reduction of various types have been demonstrated in many domains, not just at the segmental level, and not leading just to lenition (e.g. Aylett & Turk 2004, Levy & Jaeger 2007). Cohen Priva 2015 showed that low informativity is correlated with shorter segmental duration, even in domains in which lenition processes are not assumed to be frequent. Perhaps unusually short duration is a predecessor of lenition, as it is of deletion (Beckman 1996), with temporal precedence implying causation. In contrast, no accounts currently predict that reduction or omission should lead to lower information. The third alternative, in which both lenition and reduction are driven by the same unknown causes, cannot be easily ruled out, but the regression study for American English obstruent deletion, in which other factors were directly controlled for, suggests that this is not the case.

6. A POSSIBLE ALTERNATIVE: USING FUNCTIONAL-LOAD ACCOUNTS TO PREDICT LENITION. Informativity accounts use the average predictability of a single element and justify the preservation of linguistic elements based on the information they provide. FUNCTIONAL-LOAD accounts similarly argue for the preservation of contrasts based on their functional contribution to the transfer of information. Such accounts compare languages with their minimally different hypothetical counterparts in which some contrast is not preserved, and thereby estimate the relative importance of each contrast.[15]

[15] In the broadest sense of the term *functional load*, almost all information-theoretic accounts (e.g. Aylett & Turk 2006, Levy & Jaeger 2007) can be regarded as functional-load accounts, as they rely on information to justify deletion and preservation. I restrict the use of the term *functional load* to accounts that rely on the comparison between contrasts.

Functional-load accounts in this sense differ from predictability-based accounts in three major ways:

- It is not possible to evaluate the amount of information of a single linguistic element. Functional-load accounts rely on the existence or elimination of contrasts in the language. Thus, they can have widely different predictions for the elimination of the contrast between a segment and one or another minimally different segment (collapsing /t/ and /θ/ may be likely, but /t/ and /ʔ/ unlikely). Predictability-based accounts argue that some segments would be more likely to lenite, and they leave open the question of lenition's outcome. Functional-load accounts rely on both the segment and the outcome of the process.
- Functional load has no prediction for cases in which a phonological process does not lead to a collapse between contrasting segments. As far as I know, none of the authors who argued for functional load attempted to apply the approach to lenition. Many weakening processes do not collapse distinctions, and are thus not predicted by functional-load accounts.
- Collapsing two infrequent linguistic elements does not cost as much as collapsing frequent ones. This is caused by counting observed events together with unobserved events: every case in which a distinction is not lost counts toward making that distinction unnecessary. In a language in which the ratios between /t/ and /d/ and between /k/ and /g/ are both 3 : 1, and the ratio between /t/ and /k/ is 2 : 1, a functional-load account would predict that collapsing /t/ and /d/ is worse than collapsing /k/ and /g/. The predictability-based accounts presented in the previous sections, however, predict that the distinctions between /t/ and /d/ and between /k/ and /g/ could be equally important.

Andy Wedel (p.c.) argues that another important difference from predictability-based accounts is that both the minimal-pair and entropy approaches focus on the word level, whereas informativity-based approaches rely only on predictability in context, which can apply below the word level. This distinction can lead to interesting differences. The /z/ in *zebra* provides no information from contrast-based perspectives, as the rest of the word *-ebra* is unique and contrasts with no other word. Omitting the /z/ in *zebra* would not collapse any distinctions. But from the perspective of informativity-based (and other predictability-based) accounts, the /z/ would be highly informative, since few words begin with /z/.

The goal of this section is to examine how existing functional-load accounts predict segment lenition. At the segmental level, functional-load accounts predict that languages will not collapse phonemic distinctions that would hinder communication by leading speakers to confuse too many words (Martinet 1952). At least two information-theoretic measurements have been proposed to quantify functional load: differences in entropy (Hockett 1967, Surendran & Niyogi 2006) and number of minimal pairs (Martinet 1952, Wedel, Jackson, & Kaplan 2013, Wedel, Kaplan, & Jackson 2013).

6.1. Functional load as difference in entropy. The basic measurement in the quantification of functional load in Hockett 1955 and Surendran & Niyogi 2006 is the entropy of a language. In a linguistic context, the entropy of a language is the expected (mean) predictability of each linguistic element given what is already known to the listener. Consider the partial sentence in 12.

(12) An ap…

Google suggests that the user is likely to complete *an ap*… to *an apple a day*, but other completions are certainly possible. If we kept playing this game, guessing one word or one segment at a time, we would be able to estimate how predictable each word or segment is. The average predictability of all words in every context is an estimate of the entropy of English.[16] Unlike informativity, which measures the average predictability of each linguistic element separately (e.g. the informativity of /t/ vs. the informativity of /k/; informativity of *give* vs. informativity of *donate*), entropy measures the average predictability of all linguistic elements in the language (e.g. entropy of American English phonemes; entropy of American English words).

When a language distinguishes between two or more classes that are treated as identical in other languages, its entropy increases, as doing so increases the difficulty of the guessing game. For example, in a language in which there is no gender marking, it is easier to predict what the next pronoun is going to be in a context such as 13, since there is no need to distinguish between 14 and 15.

(13) No, I haven't met … yet.
(14) No, I haven't met her yet.
(15) No, I haven't met him yet.

The quantification of functional load using entropy relies on the difference in entropy between a language as it currently is and a minimally different language in which some distinction has been eliminated. The more the entropy of a language drops by eliminating a distinction, the more important that distinction is, making it unlikely that the language would lose that distinction. For example, the functional load of voicing in English can be estimated by comparing the entropy of English with the entropy of a minimally different language in which all voiced obstruents are replaced with their voiceless counterparts.

6.2. FUNCTIONAL LOAD AS NUMBER OF MINIMAL PAIRS. Wedel, Jackson, and Kaplan (2013) and Wedel, Kaplan, and Jackson (2013) contrasted several alternatives used to measure functional load, including entropy reduction and informativity. They found that the most predictive method of merger avoidance was counting the number of minimal pairs that would become homophonous if a distinction were lost. As with measuring differences in entropy, counting minimal pairs only works if distinctions are lost, which is not always the case in lenition.

6.3. APPLYING FUNCTIONAL LOAD TO LENITION PATTERNS. For cases of complete deletion, functional-load explanations do provide a prediction, since, for the most part, deletion leads to information loss (though languages may find ways to avoid the loss of information, such as compensatory lengthening; see Eckert 1985, Hochberg 1986, Gerfen 2001, Gess 2001). Consider the case of word-final deletion. If English deleted word-final /k/s, it would collapse *make* and *may*. Word-final /p/-deletion would collapse *lope* and *low*, and word-final /t/-deletion collapses *mast* and *mass*.[17] It is also possible to extend the analysis to cases in which a distinction is not completely lost, but is made perhaps less perceptually distinct.[18] Although Wedel, Kaplan, and Jackson (2013) and others did not apply their approach to lenition, I attempt to apply the approach here for the cases discussed in this article. Does functional load predict /k/-lenition in Indonesian, /t/-deletion in English, and /s/-deletion in Spanish?

[16] Shannon (1951) applies a similar character-by-character strategy to evaluate the entropy of characters in printed English.

[17] A distinction in vowel duration may remain.

[18] I am not arguing that lenition necessarily hinders perception, though several lenition processes do lead to reduced perceptibility. For instance, stops provide better place-of-articulation cues than nonstrident fricatives (Wright 2004).

First, the entropy of each language was evaluated as in 16. I then applied word-final deletions of /t/, /s/, /k/, and /p/ and measured the entropy of each minimally modified language. Differences in entropy were calculated using an unsmoothed unigram language model (Jurafsky & Martin 2009).

$$(16)\ -\sum_{word} \frac{\text{word occurrences}}{\text{all word occurrences}} \log_2 \frac{\text{word occurrences}}{\text{all word occurrences}}$$

For minimal pairs, I used the same approach, but rather than subtract the entropies of the minimally modified languages from that of the original language, I counted the number of minimal pairs added to the list of existing homophones.

For Indonesian, I used the process and corpora described in §5.2. The results are provided in Table 3.

SEGMENT	MINIMAL PAIRS	ENTROPY DIFFERENCE (IN BITS)
/s/	746	0.00471
/t/	452	0.00384
/k/	364	0.00290
/p/	126	0.00130

TABLE 3. Indonesian functional-load measurements.

For American English I used the CMU dictionary (Weide 2008) for the underlying forms of words, and combined word counts from the Fisher (Cieri et al. 2004, Cieri et al. 2005), Switchboard (Godfrey & Holliman 1997), and Buckeye (Pitt et al. 2007) corpora for word counts. The results for both functional-load measurements are in Table 4.[19]

SEGMENT	MINIMAL PAIRS	ENTROPY DIFFERENCE (IN BITS)
/s/	7,434	0.00488
/t/	5,408	0.00291
/k/	4,101	0.00170
/p/	4,058	0.00143

TABLE 4. English functional-load measurements.

For Spanish, I used the LDC Spanish CALLHOME lexicon (Garrett et al. 1996) for both counts and underlying phonological forms. Entropy differences and number of minimal pairs were calculated as for American English. The results are in Table 5.

SEGMENT	MINIMAL PAIRS	ENTROPY DIFFERENCE (IN BITS)
/s/	853	0.15689
/t/	3	0.00013
/k/	3	0.00038
/p/	0	0.00000

TABLE 5. Spanish functional-load measurements.

6.4. IMPLICATIONS FOR THE USE OF FUNCTIONAL LOAD IN LENITION. The functional-load results of American English and Spanish provide a strong bias AGAINST leniting word-final /s/ in Spanish and word-final /t/ in American English. In Spanish, deleting /s/ would yield the greatest number of minimal pairs and the largest drop in entropy. In English, deleting /t/ would lead to more minimal pairs and a greater drop in entropy than

[19] Many of the collapsed pairs are the product of an underlying *-ed* morpheme. Removing all words that end with *-ed* did not lead to qualitatively different results.

deleting /k/ or /p/, which are preserved in the same context. The Indonesian data provide some support to the functional-load account relative to /t/ and /s/, but not relative to /p/. While the entropy drop of deleting /k/ is lower than that of deleting /t/ or /s/, and fewer minimal pairs are formed by /k/ deletion than by /t/ or /s/ deletion, the opposite holds for /p/. Functional-load accounts would therefore predict /p/ to be more likely to be reduced than /k/, which is not the case. Thus, the comparative functional-load approach provided does not predict the observed lenition patterns.

It is striking that functional-load accounts predict merger avoidance but not lenition. One possibility is that different communicational pressures apply in each case. For mergers speakers may get immediate feedback when the listener confuses two words. In this case, the amount of collapsed distinctions that could lead to such confusion should have an effect on the possibility of merger, since increased confusion would lead to increased negative feedback. In contrast, Gurevich (2004) argues that lenition is often information-preserving. This is particularly true for the earlier stages of lenition (e.g. debuccalize, but do not delete; delete, but provide compensatory lengthening). Alternatively, as discussed in §5.4, lenition may rely on different mechanisms, such as reduction in duration, which is well documented for predictable contexts, both at the segmental level and at the word level.

7. SUMMARY. The actuation of sound change is neither universal nor random. Indonesian, American English, and Spanish lenite different segments while preserving segments that the other two languages lenite. This is exacerbated by English and Spanish repeatedly leniting the same segments in multiple environments and language varieties. Reviews of lenition processes (e.g. Kirchner 1998, Gurevich 2004) show this to be the case for many additional languages. This article proposes two methods that use informativity to predict lenition: crosslinguistic comparison, and controlling for phonetic factors directly in a multiple regression. No other account currently attempts to predict which language would lenite which segment, as well as the propensity to delete when phonetic factors are controlled for. Following information-preservation approaches (e.g. Cohen Priva 2012, Wedel, Kaplan, & Jackson 2013), this article uses informativity to offer a partial solution to the actuation problem and provides supporting evidence using three case studies of lenition: Indonesian /k/, American English /t/, and Spanish /s/.

Future developments of the proposed account should resolve several open questions. First, it is important to replicate the studies offered here in additional languages in order to make the results more robust. Second, this article argues that unusually low informativity licenses lenition, but it is imaginable that the relationship between the two is stronger, and low informativity actively promotes lenition. Third, this article links lenition to low informativity, but it is possible that the opposite also holds, and highly informative segments are likely to undergo fortition. Finally, it is not clear what the correct context to estimate informativity is. The current proposal takes word recognition to be the valid domain, using all previous segments in the word. Other possibilities would be to consider only phonotactic context or within-morpheme context. All of these predictions can be tested given a significantly larger set of languages.

REFERENCES

AYLETT, MATTHEW, and ALICE TURK. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47(1).31–56. DOI: 10.1177/00238309040470010201.

Aylett, Matthew, and Alice Turk. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America* 119.3048–58. DOI: 10.1121/1.2188331.

Baayen, R. Harald. 2001. *Word frequency distributions*. Dordrecht: Kluwer.

Baker, Adam; Diana Archangeli; and Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23(3).347–74. DOI: 10.1017/S0954394511000135.

Bates, Douglas; Martin Maechler; and Ben Bolker. 2014. lme4: Linear mixed-effects models using s4 classes. R package version 1.1-7. Online: http://cran.r-project.org/package=lme4.

Bauer, Laurie. 2008. Lenition revisited. *Journal of Linguistics* 44(3).605–24. DOI: 10.1017/S0022226708005331.

Beckman, Mary E. 1996. When is a syllable not a syllable? *Phonological structure and language processing: Cross-linguistic studies*, ed. by Takashi Otake and Anne Cutler, 95–124. Munich: De Gruyter.

Bell, Alan; Jason M. Brenier; Michelle Gregory; Cynthia Girand; and Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1).92–111. DOI: 10.1016/j.jml.2008.06.003.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2006. A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics* 32(2).117–66. DOI: 10.1515/TL.2006.009.

Blevins, Juliette, and Andrew Wedel. 2009. Inhibited sound change An evolutionary approach to lexical competition. *Diachronica* 26(2).143–83. DOI: 10.1075/dia.26.2.01ble.

Bresnan, Joan, and Tatiana Nikitina. 2009. The gradience of the dative alternation. *Reality exploration and discovery: Pattern interaction in language and life: Festschrift for K. P. Mohanan*, ed. by Linda Uyechi and Lian Hee Wee, 161–84. Stanford, CA: CSLI Publications.

Buck, Carl Darling. 1904. *A grammar of Oscan and Umbrian: With a collection of inscriptions and a glossary*. Boston: Ginn & Company.

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. *Usage-based models of language*, ed. by Michael Barlow and Suzanne Kemmer, 65–85. Stanford, CA: CSLI Publications.

Bybee, Joan, and Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul Hopper, 1–24. Amsterdam: John Benjamins.

Campbell, Lyle. 1996. On sound change and challenges to regularity. *The comparative method reviewed: Regularity and irregularity in language change*, ed. by Mark Durie and Malcolm Ross, 72–89. Oxford: Oxford University Press.

Campbell, Lyle. 2013. *Historical linguistics*. 3rd edn. Edinburgh: Edinburgh University Press.

Canavan, Alexandra, and George Zipperlen. 1996. CALLHOME Spanish speech. LDC96S35. Philadelphia: Linguistic Data Consortium.

Carvalho, Ana Maria. 2006. Spanish (s) aspiration as a prestige marker on the Uruguayan-Brazilian border. *Spanish in Context* 3(1).85–114. DOI: 10.1075/sic.3.1.07car.

Cieri, Christopher; David Graff; Owen Kimball; Dave Miller; and Kevin Walker. 2005. Fisher English training part 2, transcripts. LDC2005T19. Philadelphia: Linguistic Data Consortium.

Cieri, Christopher; David Miller; and Kevin Walker. 2004. The Fisher corpus: A resource for the next generations of speech-to-text. *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, 69–71. Online: http://www.lrec-conf.org/proceedings/lrec2004/pdf/767.pdf.

Clark, Urszula. 2004. The English West Midlands: Phonology. *A handbook of varieties of English, vol. 1: Phonology*, ed. by Bernd Kortmann and Edgar W. Schneider, 134–62. New York: Mouton de Gruyter.

Cohen Priva, Uriel. 2008. Using information content to predict phone deletion. *West Coast Conference on Formal Linguistics (WCCFL)* 27.90–98. Online: http://www.lingref.com/cpp/wccfl/27/abstract1820.html.

Cohen Priva, Uriel. 2012. *Sign and signal: Deriving linguistic generalizations from information utility*. Stanford, CA: Stanford University dissertation. Online: http://purl.stanford.edu/wg646gh4444.

Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology* 6(2).243–78. DOI: 10.1515/lp-2015-0008.

Daland, Robert. 2013. Variation in the input: A case study of manner class frequencies. *Journal of Child Language* 40.1091–1122. DOI: 10.1017/S0305000912000372.

de Lacy, Paul. 2002. *The formal expression of markedness*. Amherst: University of Massachusetts Amherst dissertation.

Eckert, Penelope. 1985. Grammatical constraints in phonological change: Unstressed *a in Southern France. *Orbis* 31.169–89.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Hayes et al., 232–76.

Fox, Michelle. 2001. Syllable-final /s/ lenition. LDC2001T60. Philadelphia: Linguistic Data Consortium.

Garrett, Andrew, and Keith Johnson. 2013. Phonetic bias in sound change. *Origins of sound change: Approaches to phonologization*, ed. by Alan C. L. Yu, 51–97. Oxford: Oxford University Press.

Garrett, Susan; Tom Morton; and Cynthia McLemore. 1996. CALLHOME Spanish lexicon. LDC96L16. Philadelphia: Linguistic Data Consortium.

Gerfen, Chip. 2001. A critical view of licensing by cue: Codas and obstruents in Eastern Andalusian Spanish. *Segmental phonology in optimality theory*, ed. by Linda Lombardi, 183–205. Cambridge: Cambridge University Press.

Gess, Randall. 2001. Distinctive vowel length in Old French. *Historical linguistics 1999: Selected papers from the 14th International Conference on Historical Linguistics, Vancouver, 9–13 August 1999*, ed. by Laurel J. Brinton, 145–56. Amsterdam: John Benjamins.

Godfrey, John J., and Edward Holliman. 1997. Switchboard-1 release 2. LDC97S62. Philadelphia: Linguistic Data Consortium.

Gurevich, Naomi. 2004. *Lenition and contrast: The functional consequences of certain phonetically conditioned sound changes*. New York: Routledge.

Guy, Gregory R. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3(1).1–22. DOI: 10.1017/S0954394500000429.

Hale, Mark. 2003. Neogrammarian sound change. *The handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 343–68. Cambridge, MA: Blackwell.

Han, Na-Rae; David Graff; and Myeonchul Kim. 2003. Korean telephone conversations lexicon. LDC2003L02. Philadelphia: Linguistic Data Consortium.

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

Hastie, Trevor J., and Daryl Pregibon. 1992. Generalized linear models. *Statistical models in S*, ed. by John M. Chambers and Trevor J. Hastie, 195–248. Boca Raton, FL: Chapman & Hall/CRC.

Hayes, Bruce; Robert Kirchner; and Donca Steriade (eds.) 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hickey, Raymond. 2004. Irish English: Phonology. *A handbook of varieties of English, vol. 1: Phonology*, ed. by Bernd Kortmann and Edgar W. Schneider, 68–97. New York: Mouton de Gruyter.

Hillenbrand, James M.; Michael J. Clark; and Terrance M. Nearey. 2001. Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America* 109(2).748–63. DOI: 10.1121/1.1337959.

Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3).609–21. DOI: 10.2307/415480.

Hockett, Charles F. 1955. *A manual of phonology*. Baltimore: Waverly.

Hockett, Charles F. 1965. Sound change. *Language* 41(2).185–204. DOI: 10.2307/411873.

Hockett, Charles F. 1967. The quantification of functional load. *Word* 23.300–320. DOI: 10.1080/00437956.1967.11435484.

Huang, Shudong; Xuejun Bian; Grace Wu; and Cynthia McLemore. 1996. CALLHOME Mandarin Chinese lexicon. LDC96L15. Philadelphia: Linguistic Data Consortium.

Hume, Elizabeth. 2008. Markedness and the language user. *Phonological Studies* 11.83–98.

Jaeger, T. Florian. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61(1).23–62. DOI: 10.1016/j.cogpsych.2010.02.002.

Jaeger, T. Florian, and Esteban Buz. 2017. Signal reduction and linguistic encoding. *The handbook of psycholinguistics*, ed. by Eva M. Fernández and Helen Smith Cairns. Oxford: Wiley-Blackwell, to appear.

Jurafsky, Daniel; Alan Bell; Eric Fosler-Lussier; Cynthia Girand; and William D. Raymond. 1998. Reduction of English function words in Switchboard. *Proceedings of ICSLP-98*, Sydney, 3111–14.

Jurafsky, Daniel, and James H. Martin. 2009. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Pearson Education.

Kahn, Daniel. 1976. *Syllable-based generalizations in English phonology*. Cambridge, MA: MIT dissertation.

Kaplan, Abby. 2010. *Phonology shaped by phonetics: The case of intervocalic lenition*. Santa Cruz: University of California, Santa Cruz dissertation. Online: http://roa.rutgers.edu/article/view/1107.

Kawahara, Shigeto. 2016. Japanese geminate devoicing once again: Insights from information theory. *MIT Working Papers in Linguistics (Proceedings of FAJL 8: Formal Approaches to Japanese Linguistics)* 79.43–62.

Kilany, Hanaa; H. Gadalla; Howaida Arram; A. Yacoub; Alaa El-Habashi; and Cynthia McLemore. 1997. Egyptian Colloquial Arabic lexicon. LDC99L22. Philadelphia: Linguistic Data Consortium.

Kingston, John. 2007. The phonetics-phonology interface. *The Cambridge handbook of phonology*, ed. by Paul de Lacy, 435–56. Cambridge: Cambridge University Press.

Kiparsky, Paul. 1995. The phonological basis of sound change. *The handbook of phonological theory*, ed. by John A. Goldsmith, 640–70. Cambridge, MA: Blackwell.

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

Kirchner, Robert. 2004. Consonant lenition. In Hayes et al., 313–45.

Kobayashi, Megumi; Sean Crist; Masayo Kaneko; and Cynthia McLemore. 1996. CALLHOME Japanese lexicon. LDC96L17. Philadelphia: Linguistic Data Consortium.

Kuznetsova, Alexandra; Per Bruun Brockhoff; and Rune Haubo Bojesen Christensen. 2014. lmerTest: Tests in linear mixed effects models. Online: http://CRAN.R-project.org/package=lmerTest.

Labov, William. 1965. On the mechanism of linguistic change. *Georgetown Monographs on Language and Linguistics* 18.91–114.

Labov, William. 1994. *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.

Labov, William. 2001. *Principles of linguistic change, vol. 2: Social factors*. Oxford: Blackwell.

Langstrof, Christian. 2006. Acoustic evidence for a push-chain shift in the Intermediate Period of New Zealand English. *Language Variation and Change* 18(2).141–64. DOI: 10.1017/S0954394506060078.

Lavoie, Lisa M. 2001. *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Psychology Press.

Levy, Roger, and T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. *Advances in Neural Information Processing Systems (NIPS 2006)* 19.849–56. Online: https://papers.nips.cc/paper/3129-speakers-optimize-information-density-through-syntactic-reduction.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. *Experimental phonology*, ed. by John J. Ohala and Jeri J. Jaeger, 13–44. New York: Academic Press.

Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, ed. by William J. Hardcastle and Alain Marchal, 403–39. Dordrecht: Springer.

Lindblom, Björn; Susan Guion; Susan Hura; Seung-Jae Moon; and Raquel Willerman. 1995. Is sound change adaptive? *Rivista di Linguistica* 7(1).5–36. Online: http://linguistica.sns.it/RdL/7.1/Lindblom&C.pdf.

Lindblom, Björn, and Ian Maddieson. 1988. Phonetic universals in consonant systems. *Language, speech, and mind: Studies in honour of Victoria A. Fromkin*, ed. by Larry M. Hyman and Charles N. Li, 62–78. London: Routledge.

MacKenzie, Laurel. 2010. /s/-deletion and the preservation of plurality in Modern Occitan. *University of Pennsylvania Working Papers in Linguistics* 16(2):15. Online: http://repository.upenn.edu/pwpl/vol16/iss2/15/.

Maddieson, Ian. 1984. *Patterns of sounds*. (Cambridge studies in speech science and communication.) Cambridge: Cambridge University Press.

Martinet, André. 1952. Function, structure, and sound change. *Word* 8(1).1–32. DOI: 10.1080/00437956.1952.11659416.

Mathisen, Anne Grethe. 1999. Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of language change. *Urban voices: Accent studies in the British Isles*, ed. by Paul Foulkes and Gerard J. Docherty, 107–23. London: Arnold.

Milroy, James, and Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2).339–84. DOI: 10.1017/S0022226700010306.

Moran, Steven; Daniel McCloy; and Richard Wright (eds.) 2014. *PHOIBLE online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: http://phoible.org/.

Morris, Richard E. 2000. Constraint interaction in Spanish /s/-aspiration: Three peninsular varieties. *Hispanic linguistics at the turn of the millennium: Papers from the 3rd Hispanic Linguistics Symposium*, ed. by Héctor Campos, Elena Herburger, Alfonso Morales-Front, and Thomas J. Walsh, 14–30. Somerville, MA: Cascadilla.

Nash, John C., and Ravi Varadhan. 2011. Unifying optimization algorithms to aid software system users: Optimx for R. *Journal of Statistical Software* 43(9).1–14. DOI: 10.18637/jss.v043.i09.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. *The production of speech*, ed. by Peter F. MacNeilage, 189–216. New York: Springer. DOI: 10.1007/978-1-4613-8202-7_9.

Ohala, John J. 1993a. The phonetics of sound change. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 237–78. London: Longman.

Ohala, John J. 1993b. Sound change as nature's speech perception experiment. *Speech Communication* 13.155–61. DOI: 10.1016/0167-6393(93)90067-U.

Öhman, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America* 39(1).151–68. DOI: 10.1121/1.1909864.

Palva, Heikki. 1965. *Lower Galilean Arabic: An analysis of its anaptyctic and prothetic vowels with sample texts*. Helsinki: Studia Orientalia.

Piantadosi, Steven T.; Harry J. Tily; and Edward Gibson. 2011. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America* 108(9).3526–29. DOI: 10.1073/pnas.1012551108.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul Hopper, 137–57. Amsterdam: John Benjamins.

Pitt, Mark A.; Laura Dilley; Keith Johnson; Scott Kiesling; William Raymond; Elizabeth Hume; and Eric Fosler-Lussier. 2007. Buckeye corpus of conversational speech. 2nd release. Columbus: Department of Psychology, The Ohio State University. Online: http://www.buckeyecorpus.osu.edu/.

Poplack, Shana. 1980. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. *Locating language in time and space*, ed. by William Labov, 55–67. New York: Academic Press.

R Core Team. 2014. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: http://www.R-project.org/.

Sauzet, Patrick. 2012. Occitan plurals: A case for a morpheme-based morphology. *Inflection and word formation in Romance languages*, ed. by Sascha Gaglia and Marc-Oliver Hinzelin, 179–200. Amsterdam: John Benjamins.

Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition* 133(1).140–55. DOI: 10.1016/j.cognition.2014.06.013.

Shannon, Claude E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27.379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

Shannon, Claude E. 1951. Prediction and entropy of printed English. *The Bell System Technical Journal* 30.50–64. DOI: 10.1002/j.1538-7305.1951.tb01366.x.

Sherman, D. 1975. Stop and fricative systems: A discussion of paradigmatic gaps and the question of language sampling. *Working Papers on Language Universals* 17.1–33.

Soderberg, Craig D., and Kenneth S. Olson. 2008. Indonesian. *Journal of the International Phonetic Association* 38(2).209–13. DOI: 10.1017/S0025100308003320.

Sóskuthy, Márton. 2015. Understanding change through stability: A computational study of sound change actuation. *Lingua* 163.40–60. DOI: 10.1016/j.lingua.2015.05.010.

Stevens, Kenneth N., and Arthur S. House. 1963. Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech, Language, and Hearing Research* 6.111–28. DOI: 10.1044/jshr.0602.111.

Stevens, Mary, and Jonathan Harrington. 2014. The individual and the actuation of sound change. *Loquens* 1(1):e003. DOI: 10.3989/loquens.2014.003.

Surendran, Dinoj, and Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Competing models of linguistic change: Evolution and beyond: In commemoration of Eugenio Coseriu (1921–2002)*, ed. by Ole Nedergaard Thomsen, 43–58. Amsterdam: John Benjamins.

Terrell, Tracy D. 1979. Final /s/ in Cuban Spanish. *Hispania* 62(4).599–612. DOI: 10.2307/340142.

van Son, Rob J. J. H., and Louis C. W. Pols. 2003. How efficient is speech? *Proceedings of the Institute of Phonetic Sciences* 25.171–84.

van Son, Rob J. J. H., and Jan P. H. van Santen. 2005. Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication* 47.100–123. DOI: 10.1016/j.specom.2005.06.005.

Venables, W. N., and B. D. Ripley. 2002. *Modern applied statistics with S*. 4th edn. Dordrecht: Springer.

Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wedel, Andrew; Scott Jackson; and Abby Kaplan. 2013. Functional load and the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech* 56(3).395–417. DOI: 10.1177/0023830913489096.

Wedel, Andrew; Abby Kaplan; and Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2).179–86. DOI: 10.1016/j.cognition.2013.03.002.

Weide, R. L. 2008. *The CMU pronouncing dictionary*. Release 0.7a. Pittsburgh: Carnegie Mellon University. Online: http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Weinreich, Uriel; William Labov; and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. *Directions for historical linguistics*, ed. by Winfred P. Lehmann and Yakov Malkiel, 95–195. Austin: University of Texas Press.

Wheatley, Barbara. 1996. CALLHOME Spanish transcripts. LDC96T17. Philadelphia: Linguistic Data Consortium.

Wolfe, Patricia M. 1972. *Linguistic change and the Great Vowel Shift in English*. Berkeley: University of California Press.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In Hayes et al., 34–57.

Yu, Alan C. L. 2013. Individual differences in socio-cognitive processing and the actuation of sound change. *Origins of sound change: Approaches to phonologization*, ed. by Alan C. L. Yu, 201–27. Oxford: Oxford University Press.

Zipf, George Kingsley. 1929. Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology* 15.1–95. DOI: 10.2307/310585.

ZIPF, GEORGE KINGSLEY. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Boston: Houghton Mifflin.

ZUE, VICTOR W., and MARTHA LAFERRIERE. 1979. Acoustic study of medial /t,d/ in American English. *The Journal of the Acoustical Society of America* 66(4).1039–50. DOI: 10.1121/1.383323.

Brown University
Box #1821
Providence, RI 02906
[Uriel_cohen_priva@brown.edu]

[Received 30 October 2015;
revision invited 9 May 2016;
revision received 5 August 2016;
revision invited 15 October 2016;
revision received 24 October 2016;
accepted 2 November 2016]