

See discussions, stats, and author profiles for this publication at:
<http://www.researchgate.net/publication/275773613>

On the sonority levels of fricatives and stops

ARTICLE *in* THE MENTAL LEXICON · JANUARY 2015

DOI: 10.1075/ml.10.1.04len

READS

19

2 AUTHORS, INCLUDING:



[Iris Berent](#)

Northeastern University

62 PUBLICATIONS 1,120 CITATIONS

[SEE PROFILE](#)

John Benjamins Publishing Company



This is a contribution from *The Mental Lexicon 10:1*
© 2015. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on <https://benjamins.com/#authors/rightspolicy>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

On the sonority levels of fricatives and stops

Tracy Lennertz and Iris Berent

Northeastern University

Across languages, stop-sonorant onsets are preferred to fricative-sonorant ones (e.g., *pna* > *fna*), suggesting that stop-initial onsets are better formed. Here, we ask whether this preference is active in the linguistic competence of English speakers. To address this question, we compare stop- and fricative-nasal onsets (e.g., *pnik* vs. *fnik*) to matched obstruent-obstruent controls (e.g., *ptik* vs. *fsik*, respectively). Past research has shown that (a) stop-stop onsets (e.g., *ptik*) are dispreferred to stop-nasal onsets (e.g., *pnik*); and (b) dispreferred onsets tend to be misidentified (e.g., *ptik* → *pətik*). We thus reasoned that, if fricative-nasal onsets (e.g., *fnik*) are worse formed relative to stop-nasal ones (e.g., *pnik*), then *fnik*-type onsets should be more vulnerable to misidentification, hence, their advantage over obstruent-obstruent controls (e.g., *fsik*) should be attenuated. Consequently, when compared to the obstruent-obstruent baseline (e.g., *ptik*, *fsik*), misidentification should be less prevalent in stop-nasal onsets (e.g., *pnik*) compared to fricative-nasal ones (e.g., *fnik*). The results of three experiments are consistent with this prediction. Our findings suggest that English speakers possess linguistic preferences that mirror the distribution of onset clusters across languages.

Keywords: phonology, sonority, optimality theory, syllable

English speakers intuit that *plik* could be a possible word whereas *ptik* could not, and they agree on such intuitions despite never encountering *plik* or *ptik* before (cf. Chomsky & Halle, 1965). Such observations suggest that people have productive knowledge about the sound structure of language. But what is the source of such knowledge? On one account, people's phonological knowledge is only learned by domain-general mechanisms, such as statistical learning. For example, the English preference for *plik* might reflect familiarity with *pl*-initial sequences attested in English (e.g., *play*, *plague*; Dell, Reed, Adams, & Meyer, 2000; McClelland & Plaut, 1999; Onishi, Chambers, & Fisher, 2002; Saffran, Aslin, & Newport, 1996). Sequences like *plik* may also be easier to perceive (Blevins, 2004, 2006; Ohala,

1990; Ohala & Kawasaki-Fukumori, 1997; Wright, 2004) and produce (Locke, 2000; MacNeilage & Davis, 2000; Redford, 2008) relative to sequences like *ptik*. The preference for *plik* might reflect a conjunction of its familiarity and ease of perception and production – factors that are not specific to language.

On an alternative account, people's linguistic preferences reflect broad knowledge of grammatical phonological principles that are potentially universal. Several linguistic theories hold that the grammar includes universal restrictions on sound structure (Chomsky, 1980; Prince & Smolensky, 2004). Such grammatical restrictions, for example, could disfavor syllables like *ptik* relative to *plik*. And indeed, structures like *ptik* are less frequent across languages, and any language that tolerates them tends to also allow structures like *plik* (Greenberg, 1978; cf. Berent, Steriade, Lennertz, & Vaknin, 2007). Optimality Theory (Prince & Smolensky, 2004; Smolensky & Legendre, 2006) predicts that such restrictions are active in the grammars of all speakers, even if the relevant linguistic structures (e.g., *ptik*, *plik*) are absent in their language. These predictions are indeed consistent with a large growing body of literature that has shown that people can extend their knowledge to structures that are unattested in their language (e.g., Albright, 2007; Broselow & Finer, 1991; Broselow, Chen, & Wang, 1998; Broselow & Xu, 2004; Berent et al., 2007; Berent, Lennertz, Jun, Moreno, & Smolensky, 2008; Berent, Lennertz, Smolensky, & Vaknin-Nusbaum, 2009; Daland et al., 2011; Davidson, 2006a, 2006b; Davidson, Jusczyk, & Smolensky, 2004; Jusczyk, Smolensky & Allocco, 2002; Kabak & Isardi, 2007; Moreton, 2002; 2008; Pertz & Bever, 1975; Redford, 2008; Ren, Gao, & Morgan, 2010; Wilson, 2006; Zuraw, 2007).

Our present research further investigates whether the behavior of individual speakers converge with typological regularities that are not directly attested in their language. Our case study concerns the constraints on sonority – a putative phonological property that correlates with intensity (Parker, 2002). Across languages, fricatives (e.g., *f*, *s*, *v*) are more sonorous than stops (e.g., *p*, *t*, *g*). Productive phonological restrictions in English, however, provide no evidence for the distinction between the sonority of fricatives and stops (Giegerich, 1992). Our research examines whether the behavior of English speakers is nonetheless *consistent* with a sonority distinction between stops and fricatives. The source of this consistency – whether it reflects universal grammatical constraints, phonetic restrictions or the product of inductive learning – are questions that fall beyond the scope of our present inquiry.

Sonority Restrictions on Onset Structure

The concept of sonority has been the target of much recent controversy in the phonological literature (Parker, 2012). While some authors consider sonority to be an abstract phonological property, subject to universal grammatical constraints (de Lacy, 2006), others favor a phonetic alternative (Davidson, 2010; Steriade, 2001), and underscore the role of inductive learning in acquiring sonority constraints (Daland et al., 2011; Hayes, 2011). For example, Hayes (2011) observes that typological regularities concerning sonority sequencing can be induced by a model that tracks the co-occurrence of segments in the lexicon, and is merely biased to attend to sonority. In what follows, we will outline the traditional phonological account of sonority, as this proposal forms the basis of our inquiry. Readers, however, should be mindful that the phenomena we consider are subject to alternative explanations.

Our inquiry specifically concerns the sonority restrictions on onset clusters. Onset clusters are strings of consonants occurring at the beginning of the syllable (e.g., *pl* in *please*). Although many languages allow onset clusters, not all clusters are equally preferred: onsets such as *pl* are more frequent across languages compared to onsets such as *pn*, which in turn are more frequent than onsets such as *pt*. Least frequent across languages are onsets such as *lp*. Moreover, if a language permits an infrequent onset (e.g., *lp*) it also tends to allow the more frequent ones (e.g., *pl*), but languages that allow frequent onsets do not necessarily admit less frequent ones (Greenberg, 1978, re-analyzed in Berent et al., 2007).

Several phonological theories have accounted for these facts in terms of the sonority of segments (e.g., Clements, 1990; Parker, 2002; Selkirk, 1984; Steriade, 1982). Sonority is an abstract, phonological property of sounds (Clements, 1990; Kiparsky, 1979; Steriade, 1982; Selkirk, 1984; Zec, 2007; see also de Saussure, 1916; Hooper, 1976; for discussion of its phonetic correlates, see Cassandro, Collet, Duarte, Galves, & Garcia, 2003; Galves, Garcia, Duarte, & Galves, 2002; Kawasaki-Fukumori, 1992; Ladefoged, 2001; Oudeyer, 2005; Parker, 2002; Wright, 2004). Typically, glides are the most sonorous consonants (e.g., *y*, *w*), followed by liquids (e.g., *l*, *r*), nasals (e.g., *n*, *m*), and obstruents – a group of sounds comprising both stops (e.g., *p*, *b*, *t*, *k*) and fricatives (e.g., *f*, *z*, *sh*).¹ Using these sonority levels, one can compute the sonority distance between any two consonants. For example, the onset *pl* begins with an obstruent, *p*, and rises positively in sonority to *l*, a liquid. An onset such as *pn* manifests a smaller rise, *pt* has a

1. For ease of exposition, orthography is frequently used to specify phonetic representations (e.g., */ʃ/* is transcribed as “*sh*” and */ə/* as “*e*”).

sonority plateau, it comprises two sounds from the same level, and lastly, *lp* falls in sonority. Languages vary in the minimum sonority distance that they tolerate. English requires that onsets have a large sonority rise (e.g., *pl*), whereas Russian even allows onsets with falling sonority (e.g., *lp*). But despite this cross-linguistic diversity, the sonority profile within a language is systematically constrained: if a language allows an onset with a small sonority distance (e.g., *lp*), it also tends to allow larger sonority distances (e.g., *pt*, *pn*, *pl*). In contrast, languages that allow large sonority distances do not necessarily tolerate smaller ones (data from Greenberg, 1978; reanalyzed in Berent et al., 2007). Not only can sonority capture the distribution of onsets across languages, but it can also account for syllable structure (Hooper, 1976; Prince & Smolensky, 2004; Selkirk, 1984; Smolensky, 2006; Steriade, 1982; Vennemann, 1972), syllable contact (Gouskova, 2001, 2004; Vennemann, 1972), stress assignment (de Lacy, 2007), reduplication (Morelli, 1999; Parker, 2002; Pinker & Birdsong, 1979; Steriade, 1982, 1988) and the choice of repair strategy for marked structures (Hooper, 1976).

Optimality Theory (Prince & Smolensky, 2004) attributes these typological facts to a set of universal grammatical constraints. Specifically, the *pl* > *pn* > *pt* > *lp* preference reflects a constraint that universally favors onsets with large sonority distances over small ones: a large sonority rise (e.g., *pl*) is preferred to a smaller sonority rise (e.g., *pn*), which is preferred to a sonority plateau (e.g., *pt*). Least preferred are onsets with a sonority fall (e.g., *lp*). The hypothesis that onsets with small sonority clines are universally marked does not necessarily require that sonority is explicitly encoded by the grammar. Indeed, Smolensky (2006) shows how the preference for large sonority clines can result from “independently motivated concepts” (p. 131; e.g., headed feature domains, local conjunctions, and classes of features) that do not appeal to sonority per se. Likewise, this hypothesis does not entail that the restriction on onset structure is arbitrary – a preference for large sonority distance may indeed have several phonetic explanations (Gordon, 2007; Hayes, 2004; Hayes & Steriade, 2004). Rather, the hypothesis under investigation asserts the phonological grammar includes universal constraints that render onsets with small sonority clines as marked, and the effect of these constraints is independent of the phonetic properties of the input. Crucially, these constraints are active in the grammars of all speakers, regardless of whether the particular onsets occur in one’s language. The experiments described below test whether the behavior of English speakers is consistent with this prediction.

Are Speakers Sensitive to the Sonority Distance of Onset Clusters?

A large body of research suggests that people are sensitive to the sonority distance of onset clusters that occur in their language. Attested, ill-formed onsets with small sonority distances are more difficult to produce in first-language acquisition (Barlow, 2001a, 2005; Bat-El, 2012; Gierut, 1999; Gnanadesikan, 2004; Ohala, 1999; Pater & Barlow, 2003) and they are less likely to be retained in aphasic speech (Bastiaanse, Gilbers, & van der Linde, 1994; Christman, 1992; Code & Ball, 1994; Romani & Calabrese, 1998; Stenneken, Bastiaanse, Huber, & Jacobs, 2005). Other findings suggest that sonority distance constrains performance in lexical decision tasks (Alonzo and Taft, 2002), and word games (Fowler, Treiman, & Gross, 1993; Moreton, Feng, & Smith, 2005; Treiman, 1984; Treiman, Bowey, & Bourassa, 2002; Treiman & Cassar, 1997; Treiman & Danis, 1988; Treiman & Zukowski, 1990; Yavas & Gogate, 1999).

Attested well-formed onsets, however, are typically more familiar than worse-formed attested ones, so the preference for better-formed onsets may be due to familiarity, not a grammatical restriction. Moreover, the results from attested onsets alone cannot establish the scope of sonority preferences – whether they might generalize beyond the range of sonority distances attested in speakers' own language. Fewer studies have investigated speakers' sensitivity to sonority distance using *only* unattested onset clusters. Results suggest that ill-formed unattested clusters are more difficult to accurately produce than better-formed ones (Broselow & Finer, 1991; Davidson, 2006b; Eckman & Iverson, 1993) and are judged as less likely to occur in one's native language (Albright, 2007; Coleman & Pierrehumbert, 1997; Pertz & Bever, 1975; Scholes, 1966), but whether these difficulties are specifically due to sonority remains unclear (Davidson, 2006b; Moreton, 2002).

Building on this research, Berent and colleagues (e.g., 2007, 2008, 2009) systematically examined speakers' preferences concerning the sonority distance of unattested onset clusters. In particular, Berent et al. (2007) examined English speakers' perception of highly ill-formed onsets with a sonority fall (e.g., *lp*), less ill-formed onsets with a sonority plateau (e.g., *pt*), and better-formed onsets with a small sonority rise (e.g., *pn*). People's knowledge about onset clusters was inferred from the phenomenon of misidentification.

Related work in speech perception has observed that ill-formed clusters are often misidentified to conform to native-language phonotactics (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Dupoux, Pallier, Kakehi, & Mehler, 2001; Hallé, Segui, Frauenfelder, & Meunier, 1998; Kabak & Idsardi, 2007; Massaro & Cohen, 1983; Moreton, 2002; Pitt, 1998). In particular, Pitt (1998) demonstrated that English speakers misidentify monosyllables with unattested onsets, such

as *tla*, as disyllabic (e.g., *tela*). Berent et al. (2007) suggested that these perceptual illusions might be due to the grammatical ill-formedness of such clusters. Because ill-formed onsets (e.g., *lpik*) incur a more severe violation of grammatical well-formedness constraints, they cannot be faithfully encoded by the grammar; instead, such inputs are repaired (i.e., recoded) as better-formed outputs (e.g., as *lepik*). Moreover, the probability of repair depends on the ill-formedness of the input – the worse-formed is the input, the more likely the recoding (Berent et al., 2009). So, if small sonority distances are universally ill-formed (e.g., $pn \succ pt \succ lp$) and if ill-formedness results in misidentification (e.g., $lpa \rightarrow lepa$), then the rate of misidentification should be modulated by sonority distance. That is, worse-formed monosyllables of falling sonority (e.g., *lpik*) should be more likely to be misidentified (as *lepik*) compared to better-formed monosyllables with sonority plateaus (e.g., *ptik*), and plateaus, in turn, should be more likely to be misidentified than the best-formed onsets of rising sonority (e.g., *pnik*). Crucially, speakers should be sensitive to the sonority distance of onsets that they have never heard before.

The results from numerous experiments using diverse sets of materials are consistent with this prediction (Berent et al., 2007, 2008, 2009; Berent, Balaban, Lennertz, & Vaknin-Nusbaum, 2010; Berent, Harder, & Lennertz, 2011; Berent, Lennertz, & Balaban, 2012; Berent, 2008; Berent & Lennertz, 2010). For example, in a syllable count task (e.g., “does *lpik* have one or two syllables?”), English speakers more often misidentified monosyllabic non-words of falling sonority (e.g., *lpik*) as disyllabic compared to those with sonority plateaus (e.g., *ptik*), which in turn, were more often misidentified compared to non-words of rising sonority (e.g., *pnik*). Similar results obtained when participants were explicitly asked to discriminate monosyllables from their disyllabic counterparts (e.g., “is *lpik* identical to *lepik*?”; Berent et al., 2007). Moreover, the advantage of large sonority distances obtained for different types of clusters – either obstruent- (e.g., $pna \succ pta \succ lpa$; Berent et al., 2007) or nasal-initial onsets (e.g., $ml \succ md$; Berent et al., 2009). Together, these results suggest that the rate of misidentification depends not only on the identity of the initial consonant (e.g., stops vs. nasals), but rather, their structural relation, as defined by their sonority distance.

Additional results question the possibility that the misidentification of ill-formed onsets is caused by their statistical or phonetic properties. A statistical account – the possibility that better-formed onsets (e.g., *pn*) are preferred because they resemble onsets attested in the English lexicon (e.g., *snack*) – is partly challenged by computational simulations, suggesting that the preferences of English speakers are inexplicable by models that only track the co-occurrence of segments in the English lexicon (Albright, 2007; Daland et al., 2011; Hayes, 2011). While these results do not rule out the possibility that the preferences of English

speakers might be based on co-occurrence of features (Albright, 2007; Daland et al., 2011), this explanation is further challenged by the documentation of similar preferences among speakers of Korean, a language whose lexicon lacks onset clusters altogether (Berent et al., 2008; for similar results in Mandarin, see Ren et al., 2010; Zhao & Berent, submitted, but cf., Daland et al., 2011; Duanmu, 2000; Hayes, 2011; Lee, 1994).

Other findings speak against a purely phonetic explanation. In the phonetic view, the misidentification of ill-formed onsets reflects not a grammatical repair, but rather an inability to encode the phonetic properties of such onsets from the acoustic input. However, this possibility is countered by the replication of the original findings with printed materials (Berent et al., 2009; Berent & Lennertz, 2010). Additionally, recent findings speak against the possibility that misidentification occurs because people attempt to subvocally articulate the input (and presumably fail to do so given small sonority distances), as the results replicate even when articulation is suppressed (by having participants bite on a tongue depressor; Zhao & Berent, 2013a). Taken together, the findings suggest that people share grammatical preferences regarding the sonority distance of onset clusters that do not occur in their language, and these preferences converge with the distribution of onset clusters across languages.

A New Case Study: The Sonority Levels of Fricatives and Stops

The results described so far suggest that people possess broad preferences concerning sonority *distances* that are unattested in their language (e.g., sonority plateaus vs. falls). In all these cases, however, sonority distances invariably comprised segments whose sonority levels can be discerned from active phonological restrictions in speakers' languages. For example, despite not having encountered onsets with a sonority fall (e.g., *lp*), English speakers have ample experience that could attest to the fact that the liquid *l* is more sonorous than the stop *p*. Assuming speakers are equipped with knowledge that onsets with large sonority distances are preferred to onsets with smaller ones, and given further that stop-liquid clusters (e.g., *pl*) are frequent in the lexicon, speakers can infer that the liquid *l* is more sonorous than the stop *p*. Unlike the case of stop-sonorant combinations, English does not systematically allow stop-fricative combinations, so speakers lack similar information for the ranking of stops relative to fricatives (e.g., *p* vs. *f*). These past findings do not address the question of whether knowledge of sonority *levels* (e.g., that liquids are more sonorous than stops) might be likewise universally shared. The present work examines this question. As we next demonstrate, linguistic evidence suggests that fricatives and stops differ on their sonority levels,

but the grammar of English treats them alike. Our question is whether English speakers are nonetheless sensitive to this distinction.

Various types of evidence support the possibility that fricatives are universally more sonorous than stops. Cross-linguistically, these two types of obstruents manifest different patterns of syllabification (Dell & Elmedlaoui, 1985; Hankamer & Aissen, 1974; Rose, 2000; Steriade, 1982, 1988). For example, Chaha, a language spoken in Ethiopia, distinguishes between the sonority levels of fricatives and stops at the syllable's margins. Chaha restricts coda clusters to sonority falls, all other clusters must be repaired through epenthesis. Remarkably, fricative-stop clusters are allowed (e.g., *'kift'*, "open!"), whereas stop-fricative clusters undergo epenthesis (e.g., *'kitf'* → *'kitif'*, "chop!"). Because fricative-stop clusters are treated as falls in sonority, fricatives must be represented as more sonorous than stops (Rose, 2000). Further evidence that fricatives are more sonorous than stops is present in the phonology of Attic Greek (Steriade, 1982), Pali (Hankamer & Aissen, 1974), Sanskrit (Steriade, 1988), Imdlawn Tashlhiyt Berber (Dell & Elmedlaoui, 1985), Muna (Coetzee & Pater, 2008), and Kirgiz loan-word adaptation (Gouskova, 2001). Results from first language-acquisition (Barlow, 2003; Gnanadesikan, 2004; Ohala, 1999; Pater & Barlow, 2003; Stoel-Gammon, 1985) and language-games (Barlow, 2001b) further agree with this conclusion.²

While these observations suggest a distinction between the sonority of stops and fricatives, other findings appear to challenge this conclusion. A recent set of studies by Lisa Davidson (2010, 2011) examined the production and perception of stop- and fricative-initial onset clusters in speakers of English and Catalan. Speakers produced items with an obstruent-nasal (e.g., *sm, fm, tm*) or obstruent-obstruent (e.g., *fs, fp, ps, pt*) onset; all onsets were unattested in Catalan, and all except *s*-initial ones were unattested in English. The production results indicated that speakers of both languages produced items with fricative-initial onsets more accurately than stop-initial items. Surprisingly, however, response accuracy was not further modulated by sonority distance (see also Davidson, 2000). A follow-up perceptual experiment (Davidson, 2011) likewise failed to find any effects of sonority. In that study, Catalan, English, and Russian speakers discriminated between monosyllables and their epenthetically-related disyllabic counterparts (e.g., *fpami-fepami*) in an AX identity task. As with the production study, the

2. We should also note that, among the fricatives, *s* appears to have a special status, as *s*-initial onsets present as an exception to syllable structure constraints in many languages, including English (Blevins, 1995; Selkirk, 1982), a conclusion that is further supported by experimental evidence (e.g., Gierut, 1999; Barlow, 2001b; Treiman et al., 1992). To dissociate the general question of the distinction between stops and fricatives from the status of *s*, specifically, our following experiments exclude *s*-initial onsets.

items varied as to whether they comprised an obstruent-nasal (e.g., *fn*, *tm*) or an obstruent-obstruent onset (e.g., *fs*, *fp*, *tf*). Most relevant to the current study are the results of English speakers. While English speakers tended to discriminate stop-nasal and stop-fricatives items (i.e., sonority rises; *bmada-bemada*; *pfala-pefala*) more accurately than fricative-stop items (i.e., a sonority fall; *fkada-fekada*), the results from fricative-initial items were inconsistent with the sonority hierarchy. Specifically, fricative-nasal items (i.e., sonority rises; *fnada-fenada*) elicited less accurate responses than fricative-fricative items of level sonority (e.g., *fsaga-fesaga*). Accordingly, Davidson (2011) concluded that sonority is unlikely to constrain the representation of unattested clusters.

Davidson's conclusions are unexpected given that systematic sonority effects have been documented in numerous published studies, using speakers of various language groups, stimuli types and experimental methods (Berent et al., 2007, 2008, 2009, 2010, 2011, 2012; Berent, 2008; Berent & Lennertz, 2010). We thus suspect that the divergence in outcomes is likely due to methodological factors. Indeed, Davidson's study (2011) did not systematically manipulate the sonority distance of the onset. While her fricative-initial items included both sonority rises and plateaus (e.g., *fnada*, *fsaga*), stop-initial items were comprised of sonority rises (e.g., *bmada*) – the critical stop-stop condition (i.e., plateaus) was entirely missing from her design. Similarly, each sonority-profile in her studies (2010, 2011) was represented only by a handful of items, the number of trials per condition was not balanced, and the materials varied on multiple dimensions that are unrelated to sonority, including their rhyme, voicing, compliance with the Obligatory Contour Principle (e.g., *bmada* vs. *zmafo*), and fronting. Most critically, Davidson gauged the effect of sonority by comparing onsets that differ on their initial consonant. For example, stop-nasal onsets (e.g., *bmada*, *dnapa*) were compared to fricative-nasal combinations (e.g., *fnada*, *fmatu*). A large literature suggests that fricatives and stops might present different perceptual and articulatory demands for reasons unrelated to their sonority profile. For example, it is well known that the burst release associated with stop consonants (but not fricatives) can be misidentified as an epenthetic vowel (Davidson & Shaw, 2009). In fact, variations in burst (and other phonetic) properties can explain variations in identification even when all tokens correspond to a single cluster type (e.g., distinct tokens of *bma*; cf. Davidson & Shaw, 2012; Wilson & Davidson, 2013). These observations suggest a different explanation for the null findings reported by Davidson. In this view, the failure to differentiate between the structure of fricative- and stop-initial onsets does not necessarily show that people are insensitive to sonority. Rather, genuine sonority effects could be masked by uncontrolled phonetic properties of the initial onset consonant.

The current studies revisit people's sensitivity to the sonority of fricatives and stops from a different perspective. Our critical comparison concerns the sonority profile of sonority rises and plateaus. Rather than comparing stop- and fricative-initial onsets to each other (e.g., *pnik* vs. *fnik*), we compared each of these onsets to a sonority plateau, matched for the initial consonant and rhyme (see (1)). In addition, we matched the stop- and fricative-initial onsets for their place of articulation and voicing. By comparing each of these sonority rises to a matched plateau baseline, we hoped to gauge the sonority distance of the onset while controlling for the phonetic properties of its initial consonant.

(1) The logic of our design

	Rise	Plateau baseline
Stops:	<i>pnik</i>	<i>ptik</i>
Fricatives:	<i>fnik</i>	<i>fsik</i>

If fricatives are more sonorous than stops, then the sonority rise in fricative-nasal onsets (e.g., *fn*) should be *smaller* than stop-nasal combinations (e.g., *pn*). Assuming further that onsets with small sonority distances are prone to misidentification, we further expect that, all other things being equal, misidentification should be more likely for *fn*-type onsets compared to *pn*-type onsets. Accordingly, when these rises are each compared to its respective plateau baseline (e.g., *pn* vs. *pt*; *fn* vs. *fs*), the sonority cline should be attenuated for fricative-nasal onsets compared to stop-nasal ones. Consequently, fricative-nasal onsets (e.g., *fn*) should be more frequently misidentified compared to stop-nasal ones (e.g., *pn*).

The present research tests these predictions. Our materials are monosyllabic CCVC non-words (e.g., *pnik*) and their matched CeCVC counterparts (e.g., *penik*). The critical manipulation concerns two aspects of the monosyllables: (a) the sonority distance of the onset – either small rises, plateaus or falls; and (b) the nature of the obstruent consonant in the onset – either a stop (e.g., *p*, *t*) or a fricative (e.g., *f*, *sh*). Of interest is whether people differentiate between the sonority levels of fricatives and stops. We examine this question in three experiments. In Experiments 1–2, participants determine if an auditory non-word has one syllable or two; Experiment 3 uses an identity judgment task (e.g., “is *fnik* identical to *fenik*?”). Replicating past research, we expect participants will be sensitive to sonority distance: as the sonority distance of a monosyllabic non-word decreases, people will be more likely to repair the monosyllable as a disyllable (e.g., *lpik* → *lepik*). Thus, as sonority distance decreases, people will be more likely to misidentify CCVC items as disyllabic (in Experiments 1–2) and they will misjudge them as identical to their disyllabic counterparts (in Experiment 3).

Experiment 1

Experiment 1 used the syllable-count task to examine whether English speakers distinguish between the sonority levels of fricatives and stops. In each trial, participants heard a single auditory stimulus, either a monosyllable or a disyllable (e.g., *fnik*, *fenik*), and indicated whether it included one syllable or two. Two critical aspects of the monosyllables were manipulated: the sonority distance of the onset cluster (i.e., small rise, plateau, or fall) and the type of the obstruent consonant in the onset (i.e., fricative or stop). If English speakers encode fricatives as more sonorous than stops, then the sonority rise in fricative-nasal onsets (e.g., *fn*) should be smaller, hence, fricative-nasal onsets should be more likely to be misidentified as disyllabic compared to stop-nasal ones (e.g., *pn*).

Method

Participants. Twenty-four native English speakers, undergraduate students at Northeastern University, participated in this experiment in partial fulfillment of a course requirement. In this, and all experiments, participants were considered native English speakers if they learned English at home before age 5. In this and all experiments, the majority of participants were monolinguals; about twenty-percent of participants reported exposure to another language before age 5 (e.g., Bengali, Chinese, French, Haitian Creole, Hebrew, Hindi, Korean, Polish, Portuguese, Spanish, Vietnamese).³

Materials. The experimental materials consisted of 48 pairs of non-words presented aurally. In each pair, one non-word was monosyllabic (e.g., CCVC, /fnik/) and the other was disyllabic (e.g., CəCVC, /fənik/).

All monosyllables had an unattested onset cluster that included an obstruent consonant. Two properties of the onset were manipulated: sonority distance and obstruent-type. The sonority distance manipulation contrasted three types of onsets: Onset with small rises, consisting of obstruent-nasal combinations; sonority plateaus, comprising of two obstruents (either two stops or two fricatives), and sonority falls, including a sonorant (a liquid or a nasal) followed by an obstruent. The second critical manipulation concerned the nature of the obstruent conso-

3. While some of our participants had knowledge of a second language, the syllable structure of most of those languages (e.g., Chinese, Korean) is less marked than English. In Experiment 1, the only exception was a participant reporting familiarity with Polish. The mean accuracy of this participant was 78% (the group mean was 73%).

nant in the onset – either a stop (e.g., *p, t*) or a fricative (e.g., *f, sh*). These two variables were crossed and manipulated within items. To this end, the items were arranged in triplets that were matched for their rhyme and differed only on their sonority distance (e.g., *pnik ptik, lpik*). In half of the triplets, the onset included a fricative (e.g., *f, sh*), and in the other half, the onset had a stop (e.g., *p, t*).

The monosyllabic items were subject to two sets of restrictions. First, we matched items within a triplet (i.e., small rise, plateau, fall) on several linguistic dimensions. (a) The consonants within an onset never shared the same place of articulation, as onsets with consonants sharing the same place of articulation are dispreferred across languages and may result in misidentification for reasons other than sonority distance (Kreitman, 2006). (b) Triplet members were matched for their place of articulation – in half of the triplets, onsets were labial-coronal sequences (e.g., /fnik/-/fsik/-/msik/) whereas the other half were coronal-labial sequences (e.g., /fmik/-/ffik/-/lfik/). This restriction controlled for fronting – the preference for place of articulation to move from front-to-back over back-to-front (cf. Byrd, 1992). Because English lacks labial liquids, sonority-falls beginning with a labial consonant were invariably nasal-initial (e.g., /msik/). (c) Sonority plateaus were matched for the manner of articulation – either two fricatives (e.g., /fs/) or two stops (e.g., /pt/), but *not* a fricative-stop combination (e.g., /ft/), as in our proposal, such onsets effectively exhibit a fall in sonority. (d) Triplet members were matched on voicing (e.g., sonority rises invariably comprised a voiceless-voiced sequence, cf., *fn* vs. *pn*; falls had the reverse ordering, cf., *ms* vs. *mt*); and lastly, (e) Plateaus were invariably voiceless, as sonority plateaus that disagree on voicing are dispreferred across languages (Greenberg, 1978).

Three additional aspects of our materials are noteworthy. First, to match the fricative- and stop-rises for place of articulation and voicing, we were forced to use *fm*-initial onsets for rises (the only other alternatives, /ð/ or /θ/, are both highly marked). Note, *fm* onsets are marginally attested, given their preservation in loanwords (e.g., *schmooze*) and participation in reduplication (Nevins & Vaux, 2003) – consequently some English speakers might be familiar with this combination. As we demonstrate in the General Discussion, however, our results are inconsistent with a familiarity explanation. Second, our experiments avoided voiced-stop initial onsets. Past research has shown that voiced stops manifest a burst and a voiced release, and speakers tend to misidentify these phonetic events as cues for an epenthetic vowel (Wilson & Davidson, 2013). We were concerned that the effect of voicing might attenuate the sensitivity of our experiments to sonority effects, and for this reason, our materials only used voiceless obstruents.

Finally, we adopted a set of restrictions designed to match the triplets with a fricative-onset to those with a stop-onset. First, both fricative- and stop-triplets

included only voiceless obstruents. Second, fricative-triplets and stop-triplets were matched for their place of articulation: One-half of the triplets for each obstruent-type included an onset with a labial-coronal sequence and one-half included an onset with a coronal-labial sequence. Third, each stop-triplet was matched with a fricative-triplet that had the same nucleus (i.e., vowel). Codas were selected so as to minimize the overlap with the onset consonants with respect to manner and place of articulation.

The materials also included 48 disyllabic non-words. Each such disyllable (e.g., /fənɪk/) was matched to a monosyllable (e.g., /fnɪk/) – the two were identical on all segments except for the presence of an epenthetic schwa between the two initial consonants. Likewise, the monosyllabic and disyllabic pair members were matched as closely as possible in their pitch contour and overall acoustic quality (see Figure 1 for an example).

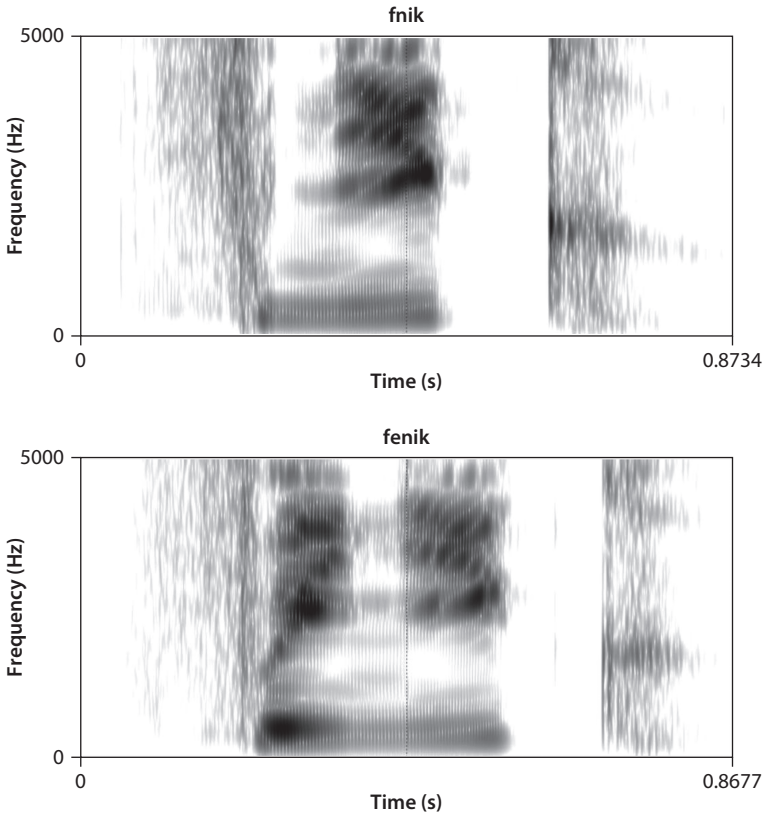


Figure 1. Spectrogram of /fnɪk/ (top) and /fənɪk/ (bottom).

In addition to the 48 pairs of experimental items described above, the experiment included 16 pairs of filler items, composed of monosyllabic non-words with a large sonority rise and their disyllabic counterparts (e.g., /flik/-/fəlik/). The monosyllabic pair member had an onset that is attested in English – an obstruent followed by a sonorant. One-half of the obstruents were fricatives (e.g., /flik/) and one-half of the obstruents were stops (e.g., /plik/). These fillers were included to encourage participants to treat the stimuli as English words. Since the perception of these items may be constrained by either their sonority distance or familiarity, they were not included in the data analyses and were not subject to the stringent linguistic controls placed on the experimental items. The Appendix lists all monosyllabic experimental and filler items.

The materials were recorded by a female native speaker of Russian. Because these onsets are all permissible in Russian, this speaker could produce them naturally. She read the items from a pseudo-randomized list presented in Cyrillic. The items were presented in the context of “X raz” (i.e., X once) and the monosyllabic pair member always preceded the disyllabic one.

Stimuli validation. In line with previous research, we predict that monosyllables with small sonority distances should be misidentified as disyllabic more often than monosyllables with larger sonority distances. Our account attributes such misidentification to the grammatical ill-formedness of small sonority distances. It is possible, however, that our auditory materials are misidentified because the Russian talker who recorded them produced articulatory artifacts that invariably signal disyllabicity. Before testing the effects of ill-formedness, we must therefore demonstrate that our materials are, in fact, identifiable as monosyllables. To this end, we presented the same materials to a group of six speakers of Russian – a language in which all these onset types are attested. If the ill-formed monosyllabic items are tainted by phonetic cues that invariably signal disyllabicity, then all listeners, irrespective of their linguistic experience, should identify these items as disyllabic. The performance of Russian participants, however (see Table 1) suggests otherwise. Russian participants identified the monosyllables with high accuracy ($M=96.18\%$). A 2 obstruent-type \times 2 sonority distance (plateaus and falls only, performance for small rises was nearly at ceiling) ANOVA yielded no significant effects with monosyllabic items (all $F < 3.51$, $p=0.10$). There were likewise no significant effects with disyllables (all $F < 1.37$, $p=0.30$). We should note that the results from a small sample should be interpreted with caution. Moreover, the possibility that Russian listeners can correctly parse the phonetic form of these inputs does not rule out the possibility that English listeners might fail to do so, an issue we revisit at the General Discussion. Nonetheless, these results do suggest that our monosyllabic items are not invariably represented as disyllabic by all listeners.

Table 1. Mean response accuracy to monosyllables in the Russian syllable count task as a function of obstruent-type and sonority distance.

	Small rise	Plateau	Fall
Fricative	100.0	97.9	93.8
Stop	93.8	97.9	93.8

Procedure. Participants, wearing headphones, were seated in front of a computer. Each trial began with a fixation point (“x”) and a message indicating the trial number. The participant pressed the space bar key to initiate the trial, triggering the presentation of a 500 ms silence, followed by the auditory stimulus. Participants were asked to determine whether the stimulus includes one syllable or two, and enter their response using the numeric keypad (1 = one syllable, 2 = two syllables). They were instructed to respond as quickly and accurately as possible. Response time was measured from the onset of the stimulus.

To familiarize participants’ with the voice of the talker, we first presented them with a brief greeting, recorded by the same talker, followed by practice trials, consisting of real English words (e.g., *drive-derive*). Feedback on response accuracy was provided in the practice trials only. Feedback on response time was not provided in either the practice or experimental trials. Each participant responded to all mono- and disyllabic items, a total of 128 experimental trials: 2 syllables (monosyllabic, disyllabic) X 2 obstruent-types (fricative, stop) X 4 sonority distances (fillers, small rise, plateau, fall) X 8 items. Trial order was randomized for each participant and the entire procedure took about 20 minutes.

Results

In this and all subsequent experiments, outliers were defined as correct responses falling 2.5 SD above the mean or faster than 200 ms and removed from the analysis of response time. We considered a response accurate if it matched the talker’s intended production (e.g., monosyllabic responses produced by the talker given monosyllabic printed inputs). In Experiment 1, outliers amounted to 3.62% of the total correct responses. We next inspected responses to monosyllabic and disyllabic items separately.

Responses to monosyllabic items. Figure 2 displays the effect of sonority distance and obstruent-type on response accuracy (the corresponding response times are provided in Table 2). In this and all figures, error bars reflect confidence intervals constructed to the difference between the mean responses for sonority distances of a given obstruent type (e.g., between stop-items with sonority rises, plateaus, and falls).

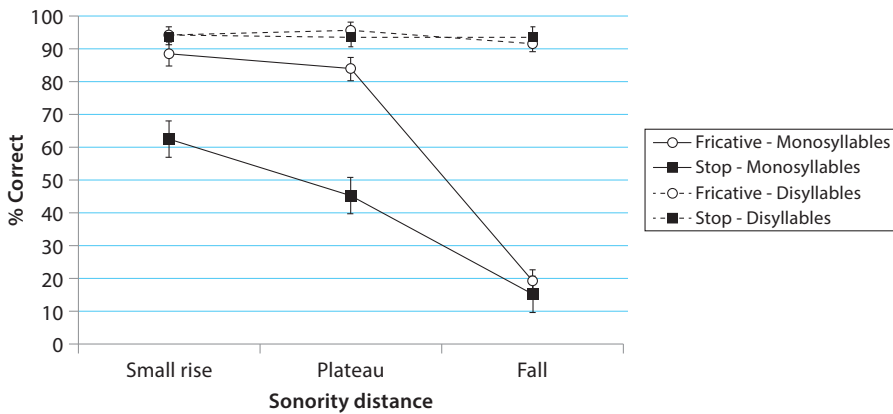


Figure 2. Mean response accuracy in Experiment 1 as a function of syllability, sonority distance, and the nature of the obstruent in the onset. Error bars represent the confidence intervals constructed for the difference among the means.

The effect of obstruent-type and sonority distance was investigated by means of 2 obstruent-type X 3 sonority distance ANOVAs conducted on response accuracy and response time using both subjects (F_1) and items (F_2) as random variables. The main effect of obstruent-type was significant in the analysis of response accuracy ($F_1(1, 23) = 57.58$, $MSE = 0.033$, $p < 0.0001$; $F_2(1, 7) = 51.87$, $MSE = 0.012$, $p < 0.0002$; in response time: both $F < 3.43$, $p = 0.10$). Similarly, sonority distance significantly modulated both response accuracy and response time (in response accuracy: $F_1(2, 46) = 145.66$, $MSE = 0.032$, $p < 0.0001$; $F_2(2, 14) = 83.71$, $MSE = 0.018$, $p < 0.0001$; in response time: $F_1(2, 16) = 24.56$, $MSE = 20792$, $p < 0.0001$; $F_2(2, 12) = 25.95$, $MSE = 17093$, $p < 0.0001$). Planned contrasts showed that participants responded significantly more accurately to onsets of rising sonority compared to sonority plateaus ($t_1(46) = 3.01$, $p < 0.005$; $t_2(14) = 2.28$, $p < 0.04$; in response time, both $p > 0.4$), which, in turn, produced more accurate ($t_1(46) = 13.05$, $p < 0.0001$; $t_2(14) = 9.89$, $p < 0.0001$) and faster ($t_1(16) = 5.57$, $p < 0.00005$; $t_2(12) = 5.75$, $p < 0.0001$) responses compared to onsets of falling sonority. Lastly, participants responded more accurately ($t_1(46) = 16.06$, $p < 0.0001$; $t_2(14) = 12.17$,

Table 2. Mean correct response time (ms) to monosyllables in Experiment 1 as a function of sonority distance and obstruent type. Standard deviations are indicated in parentheses.

	Small rise	Plateau	Fall
Fricative	1185 (140)	1211 (220)	1476 (272)
Stop	1103 (180)	1165 (125)	1435 (330)

$p < 0.0001$) and faster ($t1(16) = 6.47, p < 0.0001; t2(12) = 6.63, p < 0.0001$) to onsets of rising sonority compared to onsets of falling sonority. Crucially, the analysis of response accuracy produced a significant interaction ($F1(2, 46) = 17.78, MSE = 0.020, p < 0.0001; F2(2, 14) = 14.46, MSE = 0.008, p < 0.0004$).⁴

We next turned to examine the effect of sonority distance on response accuracy to monosyllables with stop- and fricative-onsets separately. An analysis of stop-onsets yielded a reliable simple main effect of sonority distance ($F1(2, 46) = 38.21, MSE = 0.036, p < 0.0001; F2(2, 14) = 31.44, MSE = 0.015, p < 0.0001$). Planned contrasts showed that onsets with a small sonority rise elicited more accurate responses compared to plateaus ($t1(46) = 3.13, p < .004; t2(14) = 2.84, p < 0.02$), which elicited more accurate responses compared to sonority falls ($t1(46) = 5.50, p < 0.0001; t2(14) = 4.99, p < 0.0002$).

A similar analysis of fricative-onsets also yielded a significant simple main effect of sonority distance ($F1(2, 46) = 225.42, MSE = 0.016, p < 0.0001; F2(2, 14) = 99.07, MSE = 0.012, p < 0.0001$). Like their stop-initial counterparts, fricative-initial onsets of rising and level sonority each yielded more accurate responses compared to sonority falls (*rises vs. falls*: $t1(46) = 18.99, p < 0.0001; t2(14) = 12.59, p < 0.0001$; *plateaus vs. falls*: $t1(46) = 17.71, p < 0.0001; t2(14) = 11.74, p < 0.0001$). But unlike their stop-counterparts, responses to fricative-initial onsets of rising sonority did not differ from plateaus (both $p > 0.21$). Thus, sonority rises differed from plateaus given stop, but not fricative-initial onsets.

Responses to disyllabic items. Mean response accuracy to disyllables is shown in Figure 2. A 2 obstruent-type X 3 sonority distance ANOVA on response accuracy to disyllables did not yield any significant effects (all $F < 1$). Similar ANOVAs conducted on response time yielded a reliable main effect of obstruent-type ($F1(1, 23) = 11.29, MSE = 18674, p < 0.003; F2(1, 7) = 58.37, MSE = 1115, p < 0.0002$) and a significant interaction ($F1(2, 46) = 4.21, MSE = 16909, p < 0.03; F2(2, 14) = 3.53, MSE = 8983, p < 0.058$).⁵ The effect of sonority distance was next examined separately for the disyllabic counterparts of fricative- and stop-onsets.

4. This interaction in Experiment 1 was confirmed by a mixed logit model: a 2 obstruent-type X 3 sonority distance model yielded a significant interaction ($\beta = 0.298, SE = 0.062, Z = -4.839, p < 0.0001$). As predicted, responses to onsets with rising and level sonority differed for stop ($\beta = -0.502, SE = 0.124, Z = -4.044, p < 0.0001$) but not fricative items ($\beta = -0.235, SE = 0.164, Z = -1.433, p < 0.16$). This, and all logit analyses were run using “lme4” under R, using subject and item intercepts as the random effects structure, and contrasts were set manually.

5. A mixed linear model confirmed the interaction in response time: a 2 obstruent-type X 3 sonority distance model yielded a significant interaction ($\beta = -0.019, SE = 0.005, t = -3.8, p < 0.001$).

An analysis of the disyllabic counterparts of stop-onsets yielded a simple main effect of sonority distance ($F(2, 46)=5.26$, $MSE=13105$, $p<0.009$; $F(2, 14)=4.49$, $MSE=6728$, $p<0.04$). As shown in Table 3, responses to the disyllabic counterparts of sonority falls were significantly slower compared to the counterparts of both rises ($t(46)=3.10$, $p<0.004$; $t(14)=2.78$, $p<0.02$) and plateaus ($t(46)=2.38$, $p<0.03$; $t(14)=2.36$, $p<0.04$), which, in turn, did not differ (both $p>0.5$). A similar analysis of the disyllabic counterparts of fricative-onsets did not yield a simple main effect of sonority distance (both $F<2.42$, $p=0.10$).

Table 3. Mean correct response time (ms) to disyllables in Experiment 1 as a function of sonority distance and obstruent-type. Standard deviations are indicated in parentheses.

	Small rise	Plateau	Fall
Fricative	1265 (184)	1204 (194)	1214 (219)
Stop	1109 (173)	1133 (226)	1211 (200)

Discussion

The results of Experiment 1 suggest that English speakers differentiate between fricative- and stop-initial onset clusters. Replicating previous research, stop-initial onsets of rising sonority yielded more accurate responses compared to sonority plateaus. In contrast, responses to fricative-initial onsets of rising and level sonority did not differ. People's indifference to the sonority rise in fricative-initial onsets is in line with the hypothesis that fricatives are more sonorous than stops. Because the sonority cline between an obstruent and a nasal is attenuated for fricative-initial onsets, their well-formedness is more similar to plateaus, and consequently, they are just as likely to be misidentified.

On an alternative account, the similar rate of misidentification of fricative-initial onsets with rising and level sonority may reflect a ceiling effect caused by the overall higher level of accuracy associated with the identification of fricatives. To evaluate this possibility, we divided the participants into two groups using a median split, based on their overall mean response accuracy for both mono- and disyllabic items (see Table 4; the means for the low-accuracy group and the high-accuracy group were 39% and 66%, respectively, and the difference between the means was statistically significant; $t(22)=-5.37$, $p<0.0001$). If the pattern of misidentification for fricative-initial monosyllables reflects a ceiling effect, then this pattern should be absent in the low-accuracy group. However, a separate analysis of the low-accuracy group mirrored the omnibus pattern. Specifically, the 2 obstruent-type X 3 sonority distance ANOVA yielded a significant interaction ($F(2, 22)=22.53$, $MSE=0.02$, $p<0.0001$; $F(2, 14)=21.40$, $MSE=0.01$,

$p < 0.0001$), and the simple main effect of sonority distance was significant for both fricative-initial monosyllables ($F1(2, 22) = 77.93$, $MSE = 0.02$, $p < 0.0001$; $F2(2, 14) = 53.36$, $MSE = 0.02$, $p < 0.0001$) and stop-initial ones ($F1(2, 22) = 8.81$, $MSE = 0.02$, $p < 0.002$; $F2(2, 14) = 8.12$, $MSE = 0.02$, $p < 0.005$). Planned comparisons confirmed that sonority falls produced lower accuracy than sonority rises for both fricative- ($t1(22) = 11.43$, $p < 0.0001$, $t2(14) = 9.46$, $p < 0.0001$) and stop-initial items ($t1(22) = 4.20$, $p < 0.0004$, $t2(14) = 4.03$, $p < 0.002$). However, while stop-initial items produced higher accuracy for sonority rises compared to plateaus ($t1(22) = 2.10$, $p < 0.05$, $t2(14) = 2.02$, $p < 0.064$), this contrast was not significant for fricative-initial items (both $p > 0.18$).⁶ The convergence of the low-accuracy group with the omnibus findings counters the possibility that a ceiling effect explains the results. Accordingly, the observed differences between stop- and fricative-initial onsets might suggest that English speakers distinguish between the sonority levels of fricatives and stops.

Table 4. Mean response accuracy in Experiment 1 as a function of overall accuracy, obstruent-type and sonority distance.

		Small rise	Plateau	Fall
High accuracy	Fricative	96.9	95.8	28.1
	Stop	87.5	66.7	19.8
Low accuracy	Fricative	80.2	71.9	10.4
	Stop	37.5	24.0	10.4

This proposal, however, also raises a puzzle. If fricatives are more sonorous than stops, then one might have expected sonorant-stop onsets (e.g., *lp*) with a larger fall to elicit less accurate responses compared to sonorant-fricative ones (e.g., *lf*). This, however, was not observed, possibly because this subtle distinction was overshadowed by the overwhelming propensity of sonority falls to elicit disyllabic responses. Our findings from obstruent-initial onsets indeed yield distinct patterns for stop- and fricative-initial stops. Further evidence for the great difficulty in the identification of sonority falls is seen in the responses to their disyllabic

6. This interaction for the low-accuracy group in Experiment 1 was confirmed by a mixed logit model: a 2 obstruent-type X 3 sonority distance model yielded a significant interaction ($\beta = -0.374$, $SE = 0.092$, $Z = -4.045$, $p < 0.0001$). To further investigate whether this interaction is specifically due to the stop-fricative contrast, we next tested for the simple interaction in a 2 obstruent-type X 2 sonority distance model. While that interaction did not reach significance ($\beta = 0.056$, $SE = 0.122$, $Z = 0.459$, $p < 0.65$), likely due to the small number of participants, responses to onsets with rising and level sonority differed for stop ($\beta = -0.375$, $SE = 0.169$, $Z = -2.215$, $p < 0.03$) but not fricative items ($\beta = -0.251$, $SE = 0.177$, $Z = -1.418$, $p < 0.16$), just as predicted.

counterparts. Participants were significantly slower responding to the counterparts of sonority falls (e.g., *lepik*) compared to the counterparts of sonority rises and plateaus (e.g., *penik*, *petik*). This may be a carry-over effect from their monosyllabic counterparts. Specifically, the great difficulty in the identification of monosyllables of falling sonority could have led participants to exercise caution with any sonorant-initial item, as such items are potentially monosyllables of falling sonority. Consequently, sonorant-initial items, both monosyllables and disyllables, produced slower responses compared to obstruent-initial ones.

Additional aspects of our results appear to reflect non-grammatical factors, both lexical and phonetic. Notably, fricative-initial onsets produced higher accuracy relative to their stop-initial counterparts, for both small rises ($F_1(1, 23) = 27.28$, $MSE = 0.030$, $p < 0.00003$; $F_2(1, 7) = 30.60$, $MSE = 0.009$, $p < 0.0009$) and plateaus ($F_1(1, 23) = 54.76$, $MSE = 0.033$, $p < 0.0001$; $F_2(1, 7) = 89.50$, $MSE = 0.007$, $p < 0.00004$). Because the sonority cline of fricative-initial onsets is smaller than stop-initial onsets, sonority-related factors should have only increased the ill-formedness of fricative-initial onsets, thereby increasing the propensity of such items to disyllabic repair, contrary to the observed pattern. But the lower performance with stop-initial onsets has a clear non-grammatical explanation. First, fricative-initial clusters are overall more frequent in English, and our fricative-initial stimuli were also more similar to the statistical properties of English words than stop-initial items (a detailed analysis is provided in the General Discussion, see Table 7 and Table 9).⁷ In addition, stops are more difficult to perceive than fricatives, especially when followed by another consonant. Compared to fricatives, stops contain few internal perceptual cues – many of their cues are co-articulated (e.g., formant transitions, release burst), and consequently more difficult to perceive when not adjacent to a vowel (Wright, 2004). Moreover, the speech signal of stop consonants is characterized by phonetic discontinuity – their release burst is followed by a strong reduction in the energy of all formants (Stevens, 1989). Previous research (Berent et al., 2010, cf. Davidson & Shaw, 2009) showed that participants are sensitive to such acoustic discontinuities and interpret them as cues for disyllabicity. The phonetic discontinuity of stop-initial onsets, their lack of internal perceptual cues and their lower familiarity could have all conspired to elevate their disyllabic misidentification (cf. Côté, 2000). These phonetic properties may also account for the lack of difference between stop and fricative monosyllables with an onset of falling sonority. When occurring before a vowel

7. The higher accuracy for fricative clusters might be due to the frequency of *s*-initial clusters, specifically. It turns out that this is indeed the case. Fricative-initial clusters are more frequent when *s*-initial clusters are included in the count ($t(31) = -2.10$, $p < 0.045$), but without them, stop-initial onset clusters are just as frequent as fricative-initial ones ($t(17) = 1.30$, $p > 0.20$).

(as in monosyllables with onsets of falling sonority), both stops and fricatives exhibit robust perceptual cues, resulting in similar rates of response accuracy.

Although non-phonological factors clearly affected performance in this experiment, the main finding concerns the structure of the stimuli. Not only did performance mirror the sonority profile of the onset, but it was further modulated by the status of the initial obstruent – fricative vs. stop. This finding is consistent with our hypothesis that people differentiate between the sonority levels of stops and fricatives.

Experiment 2

In Experiment 1, speakers' behavior mirrored the putatively universal distinction between the sonority levels of stops and fricatives. Replicating past research, stop-initial onsets of rising sonority elicited more accurate responses compared to stop-initial onsets comprising a sonority plateau. In contrast, fricative-initial onsets yielded no reliable differences between sonority rises and plateaus. These findings are in line with the hypothesis that English speakers consider fricatives more sonorous than stops, and consequently, the sonority rise in fricative-initial onsets is attenuated.

On an alternative account, however, the observed difference between fricatives and stops reflects not genuine differences in their sonority levels, but rather artifacts of our experimental design. Our experimental logic infers the sonority levels of fricatives and stops from their tendency to elicit disyllabic repair. Tacit in our approach is the assumption that, other things being equal, ill-formed onsets are repaired by inserting an epenthetic vowel between the two initial consonants – the greater the ill-formedness, the higher the tendency of epenthetic repair. It is conceivable, however, that fricative- and stop-initial onsets might undergo different types of repair (for similar concerns see Fleischhacker, 2001; Peperkamp, 2007): Stop-initial onsets could be repaired by inserting a schwa between the consonants (*epenthesis*, e.g., *pta* → *peta*), whereas fricative-initial onsets could be repaired by inserting a schwa before the onset cluster (*prothesis*, e.g., *fsa* → *efsa*). This possibility is worrisome because it suggests an alternative explanation for the syllable count results – an alternative we spell out below.

In this view, participants in the syllable count are sensitive to the similarity between the monosyllabic and disyllabic stimuli presented in the experiment. Essentially, participants could make their syllable count response by deciding whether a monosyllabic input (e.g., *fsik*) is distinct from the default disyllabic alternative (e.g., *fesik*). The representation of the input, however, might depend on the sonority distance of the onset (sonority plateaus are more likely to elicit repair,

irrespective of obstruent-type) and obstruent-type – stops are repaired epenthetically (e.g., *ptik* → *petik*); fricatives are repaired prothetically (e.g., *fsik* → *efsik*). But despite being equally likely to undergo repair, stop- and fricative-initial plateaus may not be equally likely to elicit a monosyllabic response. Because the encoding of stop initial-plateaus (e.g., *ptik* → *petik*) resembles the disyllables in our experiment (which are all epenthetic), they will be difficult to discriminate from disyllables hence accuracy will decline. In contrast, fricative-initial plateaus, encoded prothetically (e.g., *fsik* → *efsik*), will be quite distinct from disyllables (e.g., *fesik*), so they should be easy to discriminate from their monosyllabic counterparts – perhaps just as easy as rises. This scenario correctly predicts that sonority plateaus will produce lower accuracy than rises for stop- but not fricative-initial items despite no differences in the sonority levels of fricatives and stops. If so, the differential effects of sonority distance for fricatives and stops might occur for reasons unrelated to their sonority.

Experiment 2 was designed to evaluate this possibility. In Experiment 2, participants were once again asked to indicate whether an auditory stimulus included one syllable or two, but disyllables were now prothetically related to their monosyllabic counterparts (e.g., *fnik* vs. *efnik*). If the results of Experiment 1 are only due to use of epenthetic items for disyllables, then Experiment 2 should yield the opposite outcomes: fricative-initial onsets with rising sonority should now elicit more accurate responses than onsets with level sonority, whereas no such effect should be found with stop-initial items. On the other hand, if English speakers consider fricatives more sonorous than stops, then the pattern of Experiment 1 should replicate: fricative-initial onsets of rising and level sonority will elicit similar response accuracy, whereas stop-initial onsets of rising sonority will elicit more accurate responses than stop-initial onsets of level sonority.

Method

Participants. Sixteen native English speakers, undergraduate students at Northeastern University, participated in Experiment 2 in partial fulfillment of a course requirement. None of the participants had taken part in Experiment 1.

Materials. The materials were of the same design as Experiment 1, except disyllables included a prothetic (e.g., *efnik*) instead of an epenthetic vowel (e.g., *fenik*). The same female native Russian speaker recorded the materials. To ensure that all triplets matched in intonation and acoustic quality, both the monosyllables and disyllables were obtained from the same recording.

Procedure. The procedure was the same as Experiment 1, however the practice trials included prothetically-related real English words (e.g., *scribe-ascribe*).

Results

Correct responses falling 2.5 SD above the mean or faster than 200 ms (about 3.03% of the total observations) were excluded from the analysis of response time. Responses to monosyllabic and disyllabic items were examined separately.

The analysis of disyllabic items yielded no significant effects (in response time, all $F < 2.97$, $p = 0.08$; in response accuracy, all $F < 1.83$, $p = 0.20$). In contrast, responses to monosyllabic items were strongly modulated by both sonority distance and obstruent type (see Figure 3). An ANOVA (2 obstruent-type X 3 sonority distance) on response accuracy yielded a significant main effect of obstruent-type ($F(1, 15) = 47.29$, $MSE = 0.017$, $p < 0.0001$; $F(1, 7) = 16.91$, $MSE = 0.024$, $p < 0.005$) and sonority distance ($F(2, 30) = 168.27$, $MSE = 0.026$, $p < 0.0001$; $F(2, 14) = 112.57$, $MSE = 0.020$, $p < 0.0001$). Importantly, the interaction was significant in response accuracy ($F(2, 30) = 30.50$, $MSE = 0.014$, $p < 0.0001$; $F(2, 14) = 14.15$, $MSE = 0.015$, $p < 0.0005$).⁸ A similar analysis of response time yielded no significant effects ($F < 1.46$, $p = 0.27$).

The effect of sonority distance on response accuracy was next examined separately for monosyllables with stop- and fricative-onsets. An analysis of stop-onsets yielded a simple main effect of sonority distance ($F(2, 30) = 52.88$, $MSE = 0.026$, $p < 0.0001$; $F(2, 14) = 24.03$, $MSE = 0.028$, $p < 0.00004$). Planned contrasts demonstrated that onsets with a small rise yielded more accurate responses than plateaus ($t(30) = 3.85$, $p < 0.0006$; $t(14) = 2.60$, $p < 0.03$), which yielded more accurate responses than sonority falls ($t(30) = 6.33$, $p < 0.0001$; $t(14) = 4.27$, $p < 0.0008$).

An analysis of fricative-onsets also yielded a simple main effect of sonority distance ($F(2, 30) = 241.75$, $MSE = 0.014$, $p < 0.0001$; $F(2, 14) = 278.45$, $MSE = 0.006$, $p < 0.0001$). As in Experiment 1, fricative-initial onsets of rising and level sonority elicited more accurate responses than sonority falls (*rises vs. falls*: $t(30) = 18.66$, $p < 0.0001$; $t(14) = 20.03$, $p < 0.0001$; *plateaus vs. falls*: $t(30) = 19.40$, $p < 0.0001$; $t(14) = 20.82$, $p < 0.0001$). Crucially, however, response accuracy to fricative-initial onsets of rising sonority did not differ from plateaus (both $p > 0.4$).

8. This interaction in Experiment 2 was confirmed using a mixed logit model: a 2 obstruent-type X 3 sonority distance model yielded a significant interaction ($\beta = -0.459$, $SE = 0.084$, $Z = -5.481$, $p < 0.0001$). Additional analyses indicated that responses to onsets with rising and level sonority differed for stop- ($\beta = -0.590$, $SE = 0.147$, $Z = -4.006$, $p < 0.0001$) but not fricative-items ($\beta = 0.275$, $SE = 0.279$, $Z = 0.984$, $p < 0.33$).

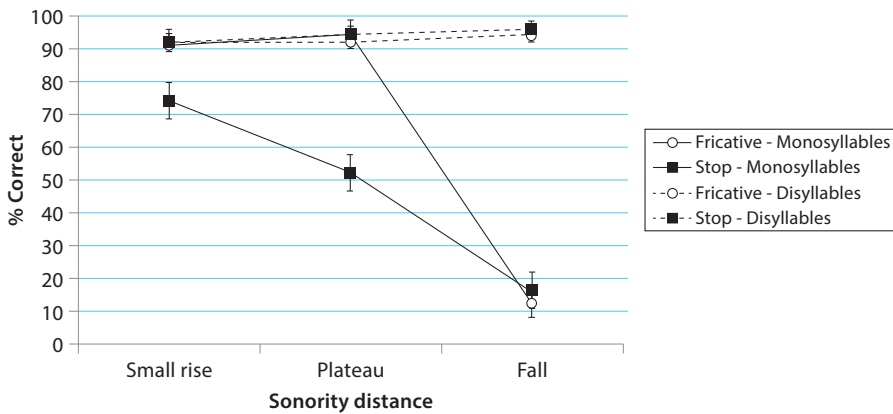


Figure 3. Mean response accuracy in Experiment 2 as a function of syllabicity, sonority distance, and the nature of the obstruent in the onset. Error bars represent confidence intervals constructed for the difference between the means.

Discussion

The results of Experiment 2 replicate the main finding of Experiment 1: participants differentiate between onsets of rising and level sonority only for stop-, but not for fricative-initial items. The persistent indifference to the sonority rise of fricative-nasal onsets, irrespective of the choice of the disyllables – with either epenthetic disyllables (e.g., *fenik* in Experiment 1) or prothetic disyllables (e.g., *efnik*; in Experiment 2) – rules out the possibility that the differences between stop- and fricative-onsets are an artifact of a particular choice of disyllables in the experiment. In fact, our findings question the contention that the disyllabic counterpart affected syllable count. Had participants performed syllable count by comparing monosyllables to their disyllabic counterparts, then the rate of monosyllabic responses to stop-initial onsets should have been higher in the presence of prothetic disyllables, as previous research (Berent et al., 2007) has shown that such items are typically repaired by epenthesis. Accordingly, an item such as *ptik* (recoded as *petik*) should be more readily discriminable from *eptik*, in Experiment 2, relative to *petik* (in Experiment 1). The comparable levels of performance across the two experiments are inconsistent with this possibility.

To further counter the possibility that the attenuated effect of sonority distance for fricatives results from a ceiling effect, we examined the performance of participants whose overall level of accuracy was low (as determined by a median split; the means for the low-accuracy group and the high-accuracy group were 51% and 63%, respectively; $t(14) = -1.80$, $p > 0.09$). The performance of these lower-accuracy participants closely mirrored the omnibus pattern. Specifically,

the 2 obstruent-type X 3 sonority distance ANOVA yielded a significant interaction ($F1(2, 14) = 12.25$, $MSE = 0.02$, $p < 0.0009$; $F2(2, 14) = 8.26$, $MSE = 0.03$, $p < 0.005$). The simple main effect of sonority distance was significant for both fricative-initial monosyllables ($F1(2, 14) = 74.56$, $MSE = 0.02$, $p < 0.0001$; $F2(2, 14) = 85.02$, $MSE = 0.02$, $p < 0.005$) and stop-initial onsets ($F1(2, 14) = 12.78$, $MSE = 0.03$, $p < 0.0007$; $F2(2, 14) = 9.90$, $MSE = 0.04$, $p < 0.003$). Planned comparisons confirmed that sonority falls ($M = 16\%$, $M = 17\%$, for the fall counterparts of fricative- vs. stop-initial items) produced lower accuracy than sonority rises for either fricative- ($M = 84\%$, $t1(14) = 10.21$, $p < 0.0001$, $t2(14) = 10.90$, $p < 0.0001$) or stop-initial items ($M = 59\%$, $t1(14) = 5.03$, $p < 0.0002$; $t2(14) = 4.42$, $p < 0.0006$). While stop-initial items tended to elicit higher accuracy for sonority rises ($M = 59\%$) compared to plateaus ($M = 42\%$, $t1(14) = 2.05$, $p < 0.06$; $t2(14) = 1.80$, $p < 0.094$), this contrast did not approach significance for fricative items (for rise: $M = 84\%$, for plateau: $M = 89\%$, both $p > 0.5$).⁹

Taken together, the results suggest that, when compared to matched plateaus, sonority rises that begin with a fricative are more vulnerable to misidentification than stop-initial counterparts. These results are consistent with the hypothesis that fricatives more sonorous than stops. The higher sonority of fricatives is expected to attenuate the sonority distance between fricatives and nasals, and diminish their advantage of plateaus.

Experiment 3

Experiments 1–2 gauged the greater ill-formedness of *fn*-type items from their propensity to elicit disyllabic responses. Experiment 3 presents a stronger test of this hypothesis. Here, people are asked to discriminate monosyllables from their disyllabic alternatives – a task that typically allows for more accurate performance than identification procedures. In each trial, participants heard two stimuli, either identical (e.g., *fnik-fnik*, *fenik-fenik*) or non-identical and epenthetically-related (e.g., *fnik-fenik*, *fenik-fnik*), and were asked to indicate whether the two items were identical. As before, two questions are of interest. First, are monosyllables with ill-formed onsets still misidentified when people are directly asked to compare them

9. The interaction for the low-accuracy group in Experiment 2 was confirmed by a mixed logit model: a 2 obstruent-type X 3 sonority distance model yielded a significant interaction ($\beta = -0.383$, $SE = 0.104$, $Z = -3.669$, $p < 0.0001$). The 2 obstruent-type X 2 sonority distance model approached significance ($\beta = 0.322$, $SE = 0.171$, $Z = 1.886$, $p < 0.06$). As predicted, responses to onsets with rising and level sonority differed for stop ($\beta = -0.426$, $SE = 0.194$, $Z = -2.194$, $p < 0.03$) but not fricative items ($\beta = 0.230$, $SE = 0.283$, $Z = 0.814$, $p > 0.42$).

to their disyllabic counterparts? Second, are responses modulated by whether the onset comprises a fricative or a stop?

We address these questions by examining people's responses to non-identical items. Generally speaking, we expect ill-formed onsets (e.g., *mtik*) to be repaired (e.g., *mtik* → *metik*), hence, harder to discriminate (e.g., from *metik*) compared to better-formed items (e.g., *pnik-penik*). But if English speakers consider fricatives more sonorous than stops, then the sonority distance between sonority rises and plateaus should be attenuated for fricative-initial monosyllables relative to their stop-initial counterparts. Consequently, the advantage of sonority rises over plateaus (i.e., faster and more accurate discrimination) should be larger for stop-compared to fricative-initial items.

Method

Participants. Thirty-six native English speakers, undergraduate students at Northeastern University, participated in Experiment 3 in partial fulfillment of a course requirement. None of the participants took part in the previous experiments.¹⁰

Materials. The materials were the same as in Experiment 1. The materials were arranged in pairs. In one-half of the trials, pair members were token-identical, either monosyllabic (e.g., *fnik-fnik*) or disyllabic (e.g., *fenik-fenik*). In the other half of the trials, the pair members were epenthetically-related (e.g., *fnik-fenik*, *fenik-fnik*). Next, the pairs were arranged in two counterbalanced lists: Each list included an equal number of identical and non-identical trials matched for sonority distance (i.e., small rise, plateau, fall), obstruent-type (i.e., fricative, stop), and presentation-order (i.e., non-identical trials were balanced for the occurrence of the monosyllabic item as either the first or second pair member). Within a list, a single item (e.g., *fnik*) was presented in either the identical trials or non-identical trials, but not in both. An additional 16 pairs of attested non-words (e.g., *flik-felik*) were included as fillers to encourage participants to treat the experimental items as English words. Each list thus comprised 128 experimental trials: 2 identity (identical, non-identical) X 2 presentation-order (the monosyllable occurred in either the first or second position) X 2 obstruent-type (fricative, stop) X 4 sonority distance (fillers, small rise, plateau, fall) X 4 items. Each participant responded to both lists, presented in counterbalanced blocks.

10. In Experiment 3, one participant reported familiarity with Hebrew (a language with a structure more marked than English). The mean accuracy of this participant was 90% (the group mean was 77%).

Procedure. Participants, wearing headphones, were seated in front of a computer. Each trial began with a fixation point (“*”) and a message indicating the trial number. The participant initiated the trial by pressing space bar, triggering the presentation of a pair of auditory stimuli, with the second stimulus following the onset of the first by 1500 ms. Participants indicated if the stimuli were identical using the numeric keypad (1 = identical, 2 = non-identical). Participants were instructed to respond quickly and were told that it was important for their response to be accurate. Response time was measured from the onset of the second stimulus.

Prior to the experimental trials, participants listened to a brief recorded greeting by the Russian talker, and completed a short practice with real English words (e.g., *drive-derive*). During the practice session, participants were provided with feedback on both accuracy and response time; in the experimental session, participants were only provided with feedback if their responses were too slow (slower than 2500 ms). Trial order was randomized. The entire procedure took about 30 minutes.

Results

Responses to identical trials were generally fast ($M = 988$ ms) and accurate ($M = 97.05\%$). Our interest is in responses to non-identical trials. Correct non-identical responses that were 2.5 SD above the mean or faster than 200 ms were excluded in the analysis of response time (about 3.02% of the total observations).

Response accuracy. Mean response accuracy to non-identical trials is shown in Figure 4. A 2 obstruent-type X 3 sonority distance ANOVA yielded a significant interaction ($F(2, 70) = 9.70$, $MSE = 0.014$, $p < 0.0002$; $F(2, 14) = 4.49$, $MSE = 0.007$, $p < 0.04$).¹¹

The simple main effect of sonority distance was significant for both stop ($F(2, 70) = 32.55$, $MSE = 0.015$, $p < 0.0001$; $F(2, 14) = 7.13$, $MSE = 0.015$, $p < 0.008$) and fricative items ($F(2, 70) = 91.53$, $MSE = 0.015$, $p < 0.0001$; $F(2, 14) = 23.58$, $MSE = 0.013$, $p < 0.00004$). Planned contrasts demonstrated that sonority rises yielded more accurate responses compared to sonority falls for either stop ($t(70) = 7.89$, $p < 0.0001$; $t(14) = 3.68$, $p < 0.003$) or fricative items ($t(70) = 13.52$, $p < 0.0001$; $t(14) = 6.87$, $p < 0.00002$). Similarly, responses to plateaus tended to be

11. A 2 obstruent-type X 3 sonority distance mixed logit model confirmed the interaction in response accuracy from Experiment 3 ($\beta = -0.147$, $SE = 0.027$, $Z = -5.432$, $p < 0.0001$). Rises elicited more accurate responses than plateaus given both fricative ($\beta = -0.649$, $SE = 0.080$, $Z = -8.08$, $p < 0.0001$) and stop items ($\beta = -0.363$, $SE = 0.064$, $Z = -5.71$, $p < 0.0001$).

more accurate compared to falls (for stop items: $t_1(70)=2.38$, $p<0.03$; t_2 , $p>0.3$; for fricative items: ($t_1(70)=6.82$, $p<0.0001$; $t_2(14)=3.46$, $p<0.004$). Crucially, however, responses to sonority rises were significantly more accurate compared to sonority plateaus with both stop ($t_1(70)=5.49$, $p<0.0001$; $t_2(14)=2.57$, $p<0.03$) and fricative items ($t_1(70)=6.70$, $p<0.0001$; $t_2(14)=3.40$, $p<0.005$). Moreover, A 2 obstruent-type X 2 sonority distance (small rises and plateaus only) ANOVA did not yield a significant interaction (both $F<1.04$, $p=0.32$). Thus, the accuracy data provide no evidence for the attenuation of the sonority rise for fricative-initial items.

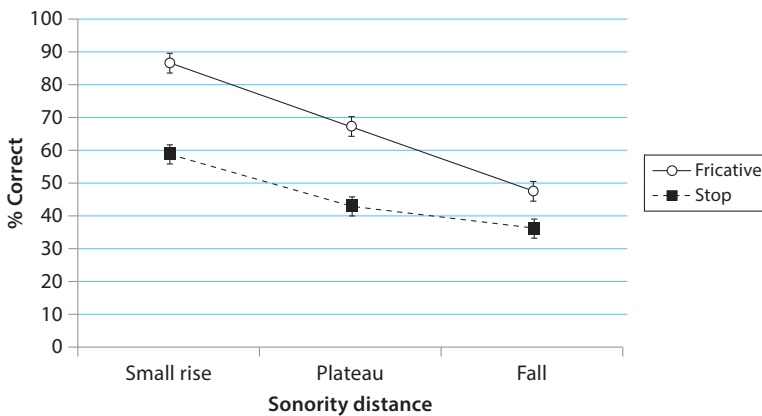


Figure 4. Mean response accuracy to non-identical trials in Experiment 3 as a function of sonority distance and the nature of the obstruent in the onset. Error bars reflect confidence intervals constructed for the difference between the means.

Response time. Mean correct response time for non-identical trials as a function of sonority distance and obstruent-type is presented in Figure 5. An inspection of the means suggests that the difference in sonority distance between sonority rises and plateaus was smaller for fricative-initial onsets compared to stop-initial ones. These conclusions are supported by the significance of the interaction in a 2 obstruent-type X 3 sonority distance ANOVA ($F_1(2, 68)=6.19$, $MSE=6699$, $p<0.004$; $F_2(2, 14)=5.45$, $MSE=1444$, $p<0.02$).¹²

12. A 2 obstruent-type X 3 sonority distance mixed linear model confirmed the interaction in response time in Experiment 3 ($\beta=-0.015$, $SE=0.004$, $t=-3.8$ $p<0.0002$). The 2 obstruent-type X 2 sonority distance (small rises and plateaus only) model in response time also yielded a significant interaction ($\beta=-0.012$, $SE=0.006$, $t=-2.1$ $p<0.038$).

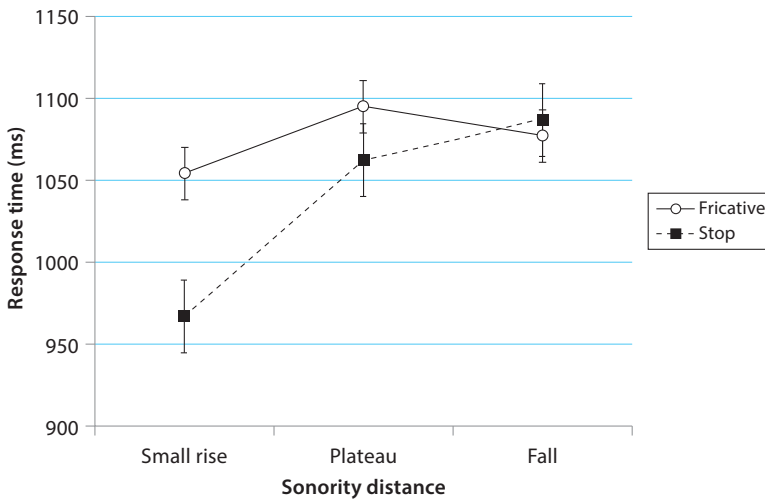


Figure 5. Mean correct response time for non-identical trials in Experiment 3 as a function of sonority distance and the nature of the obstruent in the onset. Error bars reflect confidence intervals constructed for the difference between the means.

Responses to non-identical trials with stop- and fricative-onsets were next examined separately. The analysis of fricative items did not yield a reliable simple effect of sonority distance (both $F < 2.54$, $p = 0.09$). In contrast, sonority distance reliably modulated responses to stop items ($F(2, 68) = 16.82$, $MSE = 8362$, $p < 0.0001$; $F(2, 14) = 20.42$, $MSE = 1665$, $p < 0.0001$). Planned contrasts showed that onsets of falling sonority yielded reliably slower responses compared to sonority rises ($t(68) = 5.49$, $p < 0.0001$; $t(14) = 6.39$, $p < 0.00003$), and marginally so compared to plateaus ($t(1, p > 0.26$; $t(14) = 3.07$, $p < 0.009$). Crucially, however, responses to sonority plateaus were also slower than sonority rises ($t(68) = 4.36$, $p < 0.0001$; $t(14) = 3.32$, $p < 0.006$).

While these results are consistent with our predictions, one might worry that these differences in response time might result from extraneous sources. Recall that our experiments measured response time from the onset of the second item, so one might be concerned that the observed differences in response time could reflect differences in stimuli duration, not sonority effects per se. To investigate this possibility, we assessed the unique effects of sonority distance and duration by means of two linear regression analyses conducted on responses to sonority rises and plateaus. To determine the unique effect of sonority distance, we first forced the target's duration and syllable structure (monosyllable vs. disyllable) into the model, whereas sonority distance was entered last. The effect of stimulus duration was assessed by reversing the order of the predictors (sonority distance

and syllable structure were entered first; duration was forced last). Results showed that stimulus duration (entered as the last predictor) did not systematically affect responses – its effect was significant only for fricatives ($R^2 = .255$, $F(1, 28) = 12.89$, $p < 0.001$), but not for stops ($R^2 = .010$, $F < 1$, $p = 0.52$). In contrast, the effect of sonority remained significant for stop-initial items after controlling for stimulus duration ($R^2 = .237$, $F(1, 28) = 10.20$, $p < 0.004$). Although the unique contribution of sonority distance was now significant also for fricative-initial items ($R^2 = .103$, $F(1, 28) = 5.22$, $p < 0.04$), this effect accounted for less than half of the variance explained by the sonority of stop-initial items. These analyses suggest that our conclusions cannot be reduced to differences in length between stimuli. After controlling for stimulus duration, the effect of sonority distance remains attenuated for fricatives relative to stops.

Discussion

Experiment 3 examined the representation of sonority distance using a discrimination task. Replicating previous findings (Berent et al., 2007), performance was sensitive to the sonority manipulation: As sonority distance decreased, people were less accurate and slower to distinguish monosyllables from their disyllabic counterparts. Moreover, this effect was further modulated by the type of obstruent – stop or fricative. While stop-initial onsets of rising sonority elicited faster responses than their level-sonority counterparts, the rise in sonority conferred a smaller advantage to fricative-initial onsets.¹³ These results are in line with the hypothesis that native English speakers differentiate between the unattested sonority levels of stops and fricatives.

13. Unlike syllable count, in the present experiment, obstruent-type did not modulate the effect of sonority distance in response accuracy. This discrepancy might well be due to the change in task demands. The explicit comparison of monosyllables with their disyllabic counterparts could have allowed participants to discriminate among them even when the sonority rise was attenuated for fricative-initial onsets. Doing so, however, incurred a cost in response time. Indeed, responses to fricative-initial onsets were overall slower than responses to stop-initial onsets. Moreover, the advantage of sonority rises compared to plateaus was smaller for fricative-initial onsets compared to stop-initial ones.

General Discussion

The present research investigated whether speakers are sensitive to putatively universal restrictions on the sound structure of language. Our specific case concerned restrictions on the sonority levels of fricatives and stops. While English phonotactics (e.g., the restrictions on attested onsets and phonological alternations) do not distinguish between fricatives and stops (Giegerich, 1992), this distinction is present in several other languages. We examined whether English speakers are nonetheless sensitive to this difference between the sonority levels of fricatives and stops.

If English speakers consider fricatives more sonorous than stops, then the sonority distance between small rises and plateaus should be attenuated for fricative-initial onsets (e.g., *fn* vs. *fs*) compared to stop-initial ones (e.g., *pn* vs. *pt*). Consequently, the sonority rise in a stop-nasal onset (e.g., *pn*) should be larger than the rise in a fricative-nasal onset (e.g., *fn*). Given that ill-formed onsets with small sonority distances are more likely to be repaired as disyllabic (Berent et al., 2007, 2008, 2009), we predicted that performance should vary depending on the status of the initial obstruent as a stop or fricative. Specifically, compared to their matched plateaus (e.g., *fs*, *pt*), fricative-initial monosyllables with rising sonority (e.g., *fn*) should exhibit a lower rate of repair than stop-initial ones (e.g., *pn*). Our results are consistent with this prediction. In each experiment, sonority distance was modulated by the status of the obstruent – fricative or stop. In Experiments 1 and 2, both syllable count tasks, the rate of misidentification differed between stop-initial and fricative-initial monosyllables. In line with previous research, stop-initial monosyllables with a small rise (e.g., *pna*) elicited more accurate responses than stop-initial monosyllables with a plateau (e.g., *pta*). In contrast, the identification of fricative-initial monosyllables with a small rise (e.g., *fna*) and plateau (e.g., *fsa*) did not reliably differ. Experiment 3, an identity judgment task, provided further evidence for the distinction between fricatives and stops: Sonority rises produced faster responses than plateaus for items comprising stop-initial onsets, but this effect was not systematically found with fricative-initial onsets. Taken together, our results suggest that English speakers failed to differentiate between fricative-initial monosyllables of rising and level sonority (e.g., *fna* vs. *fsa*). In contrast, stop-initial monosyllables of rising sonority consistently yielded superior identification relative to sonority plateaus (e.g., *pna* vs. *pta*).

While this pattern follows straightforwardly from the sonority hypothesis, on an alternative lexical account, these findings could conceivably result from the similarity of our materials to the onsets attested in English. Accordingly, the knowledge guiding participants' behavior could reflect the statistical properties of the English lexicon. We consider several such lexical statistical accounts below.

The Role of Attestation

One explanation is presented by the possibility that some of the onsets employed in our experiments are, in fact, attested in the lexicon of our participants. Specifically, one-half of the fricative-initial onsets with a small rise began with “*shm*” – an onset cluster that is marginally attested in English, as it is preserved in loanwords (e.g., *schmooze*), and takes part in productive reduplication (Nevins & Vaux, 2003, e.g., *homework-shmomework*). As explained earlier (under Method, Section 2.1), several considerations have led us to favor *shm*-onsets as the best exponent of their category (the category of voiceless coronal obstruents). One might worry, however, that the identification of these onsets might reflect their attestation, rather than sonority profile.

Thankfully, the sonority and attestation accounts make different predictions that allow us to adjudicate between these explanations. On the sonority account, the identification of “*shma*” depends on its sonority profile. Since *shma*-type items have a smaller sonority cline than stop rises (e.g., *tma*), *shma*-items should be more prone to misidentification. Accordingly, when *shma*- and *tma*-type items are each compared with their plateau baselines (e.g., *shfa* and *tpa*, respectively), the advantage of *shma*-rise should be attenuated relative the *tma*-rise. Moreover, since the sonority cline in *shma* is comparable to *fna*, a sonority account should likewise predict that these items should produce similar results when compared to their respective plateau controls (e.g., *shfa* and *fsa*, respectively). The predictions of the attestation account are quite different. In this view, the attested *shma* onsets are preferred to unattested onsets, so the identification of *shma* should be invariably superior to unattested rises – either *fna* or *tma*.

To test these predictions, we compared the identification of *shma*- and *tma*-type onsets to their plateau controls. The means for the relevant dependent measures are shown in Table 5. As shown in Table 6, in no case did a 2 place (labial vs. coronal) X 2 sonority distance (small rise vs. plateaus) ANOVA yield a significant interaction. While responses to coronal fricatives are overall more accurate (Experiments 1–2) and faster (Experiment 3) than labial fricatives, the effect of sonority is similar across the groups: the sonority cline for fricative-nasal onsets is attenuated for both places of articulation (e.g., *shma-shfa* and *fna-fsa*). These results are inconsistent with the attestation explanation, but they are fully in line with the sonority account.

Table 5. Response accuracy (Experiments 1 and 2) and correct response time (Experiment 3) by sonority distance for onsets with a coronal- or labial-onset.

			Small rise	Plateau	Difference
Experiment 1	Coronal	Fricative	96.88	87.5	9.38
		Stop	64.58	45.83	18.75
	Labial	Fricative	80.21	80.21	0
		Stop	60.42	44.79	15.63
Experiment 2	Coronal	Fricative	95.31	93.75	1.56
		Stop	68.75	53.13	15.62
	Labial	Fricative	87.5	95.31	-7.81
		Stop	79.69	51.56	28.13
Experiment 3	Coronal	Fricative	1030	1075	-45
		Stop	952	1037	-85
	Labial	Fricative	1086	1105	-19
		Stop	973	1062	-89

Table 6. The resulting interactions from a 2 place (labial vs. coronal) X 2 sonority distance (small rise vs. plateau) ANOVA by Experiment.

Experiment	Random factor	Fricatives				Stops			
		MS	df	F	p-value	MS	df	F	p-value
1	subjects	.053	1, 23	1.12	NS	.006	1, 23	<1	NS
	items	.009	1, 6	2.93	NS	.001	1, 6	<1	NS
2	subjects	.035	1, 15	1.90	NS	.063	1, 15	1.03	NS
	items	.009	1, 6	2.35	NS	.016	1, 6	<1	NS
3	subjects	5495	1, 34	1.08	NS	77	1, 31	<1	NS
	items	123	1, 6	<1	NS	1	1, 6	<1	NS

Segment Co-occurrence

Another statistical account attributes our findings to the co-occurrence of segments in the English lexicon (e.g., Bailey & Hahn, 2001; Coleman & Pierrehumbert, 1997; Frisch, Large, & Pisoni, 2000; Hay, Pierrehumbert, & Beckman, 2003; Vitevitch & Luce, 2004; Sibley, Kello, Plaut, & Elman, 2008; for some experimental evidence, see Friederici & Wessels, 1993; Hay, Pelucchi, Graf Estes, & Saffran, 2011; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Pelucchi, Hay, & Saffran, 2009; Saffran et al., 1996; Saffran, 2003). Participants might nonetheless possess tacit statistical knowledge concerning the co-occurrence of their segments that could inform our pattern of results. In this view, stop-initial monosyllables of level sonority may be misidentified because their statistical properties

are less frequent than sonority rises (e.g., greater familiarity with *pna* than with *pta*). In contrast, fricative-initial onsets of rising and level sonority elicit similar behavior because their statistical properties are equally unfamiliar (e.g., *fna* = *fsa*).

To test this explanation, we first examined the statistical properties of our monosyllabic items using several measures of segment co-occurrence. For each of our items, we calculated the number of neighbors (i.e., words created by adding, substituting, or deleting a single phoneme), neighbors' frequency (i.e., summed frequency of neighbors), the position-specific phoneme probability (i.e., the probability that a phoneme occurs in a given position, mean across the four positions) and bi-phone probability (i.e., the probability that a pair of adjacent phonemes occurs in a given position, mean across the three positions).¹⁴ As shown in Table 7, these statistical patterns do not match our results. First, *pn*-type items with a small rise are not invariably more frequent than *pt*-type ones – on the contrary, *pn*-type items actually tend to be less frequent than *pt*-type members. Second, the difference between monosyllables with rising and level sonority is not consistently larger for stop-initial items compared to fricative-initial ones.

Table 7. Statistical properties of the fricative- and stop-initial monosyllables. The averaged values across 8 items per onset-type are shown.

	Fricatives		Stops	
	Small rise	Plateau	Small rise	Plateau
Number of neighbors	0.75	0.625	1.125	1.75
Neighbors' frequency (summed)	18.38	13.88	9.50	30.38
Position-specific phoneme probability	0.0351	0.0273	0.0392	0.0354
Bi-phone probability	0.0006	0.0005	0.0006	0.0007

To further test this statistical account, we submitted the data from each experiment to a stepwise linear regression, controlling for the statistical properties of the monosyllable in the first step and examining the unique effect of sonority distance in the second step. We also assessed the unique contribution of statistical properties by entering them last, after controlling for sonority distance in the first step. Since we are specifically interested in accounting for the *pn* > *pt* and *fn* = *fs* pattern, we only consider sonority rises and plateaus. The results are shown in Table 8.

14. Phoneme and bi-phone probabilities were obtained from the Phonotactic Probability Calculator prepared by Vitevitch and Luce (2004). Neighborhood properties were obtained from the Speech and Hearing Lab database at <http://128.252.27.56/Neighborhood/Home.asp>. When accessed (November 2009) this database failed to distinguish between the letter case of the phonological input (e.g., “S” corresponding to /ʃ/ and “s” corresponding to /s/). Two researchers manually inspected the output for its accuracy.

As expected, sonority distance had no unique effect on the identification of fricative-initial onsets. In contrast, sonority distance uniquely accounted for the identification of stop-initial onsets even after controlling for their statistical properties in the initial step. In fact, statistical familiarity did not capture the results in any of our experiments. In contrast, the effect of sonority distance was always significant for each experiment.

Table 8. The results of a series of step-wise linear regressions from Experiments 1–3.

Experiment	Last predictor	Fricatives				Stops			
		R ² change	df	F	p-value	R ² change	Df	F	p-value
1	sonority distance	.096	1, 10	2.10	NS	.364	1, 10	7.57	.020
	statistical properties	.490	4, 10	2.67	NS	.129	4, 10	< 1	NS
2	sonority distance	.005	1, 10	< 1	NS	.186	1, 10	4.62	.057
	statistical properties	.300	4, 10	1.19	NS	.296	4, 10	1.83	NS
3	sonority distance	.275	1, 10	5.57	.04	.535	1, 10	12.77	.005
	statistical properties	.346	4, 10	1.76	NS	.169	4, 10	1.01	NS

Feature Co-occurrence

The findings reviewed so far suggest that the distinction between stop- and fricative-initial onsets is inexplicable by the co-occurrence of their segments in the English lexicon. It is conceivable, however, that that distinction could reflect the statistical co-occurrence of features – a possibility that has received recent support (Adriaans & Kager, 2010; Albright, 2009; Daland et al., 2011; Frisch, Pierrehumbert, & Broe, 2004; Hayes, 2011; Hayes & Wilson, 2008;). A preliminary inspection of the English lexicon, however, reveals some discrepancies with our findings.

For example, consider the co-occurrence of the manner feature in English onsets. Although at first glance, the number of attested fricative- and stop-initial onset clusters in English (see Table 9) appears to match our results (i.e., fricatives co-occur with both sonorants and obstruents, whereas stops only occur with sonorants), a closer inspection suggests otherwise. Our results show comparable

outcomes for fricative-nasal and fricative-fricative onsets, but the former type (e.g., *snail*, *small*) is far more frequent than the latter (the only fricative-fricative onset in English is /sf/, e.g., *sphere*).

Another discrepancy is evident with respect to the co-occurrence of place of articulation. All fricative-nasal onsets attested in English are coronal – no such onsets begin with a labial fricative – so an inductive learner is likely to show different patterns for labial-fricative (e.g., /f/) and coronal-fricative (e.g., /ʃ/) items. That is, the attenuated contrast between sonority rises and plateaus is only expected for coronal-initial fricatives (i.e., $f_m = ff$); labial-initial fricatives should pattern with stops (i.e., an advantage for f_n relative to f_s). English participants, however, are oblivious to this fact. A 2 place of articulation (labial vs. coronal) by 2 sonority distance (small rise vs. plateau) ANOVA on the relevant responses from each experiment found no hint of an interaction (see Tables 5 and 6). While a full assessment of a statistical account calls for a detailed computational analysis – a task that falls beyond the scope of our present paper, the available evidence suggests that the induction of the relevant knowledge from the lexicon might not be trivial.

Table 9. The number of attested fricative- and stop-initial onset clusters by manner of articulation (e.g., sonorant or obstruent).

Type of cluster	Fricative-initial	Stop-initial
Obstruent-sonorant (e.g., <i>fl</i> , <i>pl</i>)	9	14
Obstruent-obstruent (e.g., <i>sp</i>)	10	0

Conclusion

Across languages, fricatives are more sonorous than stops. Our experimental results are consistent with the possibility that English speakers likewise encode this sonority distinction. Regression analyses challenge the attribution of these results purely to the statistical properties of the English lexicon. While our findings cannot fully reject such statistical explanations, the conclusions open up the possibility that English speakers possess algebraic grammatical knowledge that renders fricatives more sonorous than stops (cf. Berent, 2013). Our results do not necessarily show that this ranking is universal or experience independent. It is in fact likely that participants can infer the sonority hierarchy from experience with the phonetic properties of fricatives and stops (cf. Hayes & Steriade, 2004; Parker, 2002, 2008; Wright, 2004). Nonetheless, grammatical restrictions on sonority appear to be central to the language system, as they are evident across language

modalities – spoken and signed (Brentari, 1993, 1998; Padden & Perlmutter, 1987; Perlmutter, 1992; Sandler, 1993; Sandler and Lillo-Martin, 2006). Moreover, speakers spontaneously transfer sonority restrictions on the syllable from speech to sign (Berent, Dupuis, & Brentari, 2013). The inferential process that allows speakers to extract this information, and the mechanisms that compel the grammars of disparate languages to converge on their design await future research.

References

- Adriaans, F., & Kager, R. (2010). Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language*, 62, 311–331. DOI: 10.1016/j.jml.2009.11.007
- Albright, A. (2007). *Natural classes are not enough: biased generalization in novel onset clusters*. Manuscript, Cambridge, MA: MIT.
- Albright, A. (2009). Feature-based generalisation as a source of gradient acceptability. *Phonology*, 26, 9–41. DOI: 10.1017/S0952675709001705
- Alonzo, A., & Taft, M. (2002). Sonority constraints on onset-rime cohesion: Evidence from native and bilingual Filipino readers of English. *Brain and Language*, 81, 368–383. DOI: 10.1006/brln.2001.2531
- Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44, 568–591. DOI: 10.1006/jmla.2000.2756
- Barlow, J. A. (2001a). A preliminary typology of initial clusters in acquisition. *Clinical Linguistics & Phonetics*, 15, 9–13. DOI: 10.3109/02699200109167622
- Barlow, J. A. (2001b). Individual differences in the production of initial consonant sequences in Pig Latin. *Lingua*, 111, 667–696. DOI: 10.1016/S0024-3841(00)00043-7
- Barlow, J. A. (2003). Asymmetries in the acquisition of consonant clusters in Spanish. *Canadian Journal of Linguistics*, 48, 179–210. DOI: 10.1353/cjl.2004.0024
- Barlow, J. A. (2005). Sonority effects in the production of consonant clusters by Spanish-speaking children. In D. Eddington (Ed.) *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 1–14). Somerville, MA: Cascadilla Proceedings Project.
- Bastiaanse, R., Gilbers, D., & van der Linde, K. (1994). Sonority substitutions in Broca's and conduction aphasia. *Journal of Neurolinguistics*, 8, 247–255. DOI: 10.1016/0911-6044(94)90011-6
- Bat-El, O. (2012). The Sonority Dispersion Principle in the acquisition of Hebrew word final codas. In S. Parker (Ed.), *The sonority controversy* (pp. 319–344). Berlin: Mouton de Gruyter.
- Berent, I. (2008). Are phonological representations of printed and spoken language isomorphic? Evidence from the restrictions on unattested onsets. *Journal of Experimental Psychology: Human Perception & Performance*, 34, 1288–1304. DOI: 10.1037/0096-1523.34.5.1288
- Berent, I. (2013). *The phonological mind*. Cambridge, UK: Cambridge University Press.

- Berent, I., Balaban, E., Lennertz, T., & Vaknin-Nusbaum, V. (2010). Phonological universals constrain the processing of nonspeech stimuli. *Journal of Experimental Psychology: General*, 139, 418–435. DOI: 10.1037/a0020094
- Berent, I., Dupuis, A., & Brentari, D. (2013). Amodal aspects of linguistic design. *PLoS ONE*, DOI: 10.1371/journal.pone.0060617.
- Berent, I., Harder, K., & Lennertz, T. (2011). Phonological universals in early childhood: Evidence from sonority restrictions. *Language Acquisition*, 18, 281–293. DOI: 10.1080/10489223.2011.580676
- Berent, I., & Lennertz, T. (2010). Universal constraints on the sound structure of language: Phonological or acoustic? *Journal of Experimental Psychology: Human Perception & Performance*, 36, 212–223. DOI: 10.1037/a0017638
- Berent, I., Lennertz, T., & Balaban, E. (2012). Language universals and misidentification: A two-way street. *Language and Speech*, 55, 311–330. DOI: 10.1177/0023830911417804
- Berent, I., Lennertz, T., Jun, J., Moreno, M. A., & Smolensky, P. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences*, 105, 5321–5325. DOI: 10.1073/pnas.0801469105
- Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, 26, 75–108. DOI: 10.1017/S0952675709001729
- Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104, 591–630. DOI: 10.1016/j.cognition.2006.05.015
- Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Cambridge, MA: Blackwell Publishers.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511486357
- Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics*, 32, 117–166.
- Brentari, D. (1993). Establishing a sonority hierarchy in American Sign Language: The use of simultaneous structure in phonology. *Phonology*, 10, 281–306. DOI: 10.1017/S0952675700000063
- Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.
- Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, 7, 35–59. DOI: 10.1177/026765839100700102
- Broselow, E., Chen, S-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20, 261–280. DOI: 10.1017/S0272263198002071
- Broselow, E., & Xu, Z. (2004). Differential difficulty in the acquisition of second language phonology. *International Journal of English Studies*, 4, 135–163.
- Byrd, D. (1992). Perception of assimilation in consonant clusters: A gestural model. *Phonetica*, 49, 1–24. DOI: 10.1159/000261900
- Cassandro, M., Collet, P., Duarte, D., Galves, A., & Garcia, J. (2003). *A universal linear relation among acoustic correlates of rhythm*. Unpublished manuscript.
- Chomsky, N. (1980). *Rules and representations*. Columbia University Press: New York.

- Chomsky, N., & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of Linguistics*, 1, 97–138. DOI: 10.1017/S0022226700001134
- Christman, S.S. (1992). Uncovering phonological regularity in neologisms: contributions of sonority theory. *Clinical Linguistics & Phonetics*, 6, 219–247. DOI: 10.3109/02699209208985532
- Clements, G.N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech* (pp. 282–333). Cambridge, UK: Cambridge University Press.
- Code, C., & Ball, M.J. (1994). Syllabification in aphasic recurring utterances: Contributions of sonority theory. *Journal of Neurolinguistics*, 8, 257–265. DOI: 10.1016/0911-6044(94)90012-4
- Coleman, J., & Pierrehumbert, J.B. (1997). Stochastic phonological grammars and acceptability. In J. Coleman (Ed.), *Proceedings of the 3rd Meeting of the ACL Special Interest Group in Computational Phonology* (pp. 49–56). Somerset, NJ: Association for Computational Linguistics.
- Coetzee, A.W., & Pater, J. (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna. *Natural language and linguistic theory*, 26, 289–337. DOI: 10.1007/s11049-008-9039-z
- Côté, M.-H. (2000). *Consonant cluster phonotactics: A perceptual approach*. Doctoral dissertation, MIT Working Papers in Linguistics.
- Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology*, 28, 197–234. DOI: 10.1017/S0952675711000145
- Davidson, L. (2000). Experimentally uncovering hidden strata in English phonology. In L. Gleitman & A. Joshi (Eds.), *Proceedings of the 22nd annual conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Davidson, L., Jusczyk, P.W., & Smolensky, P. (2004). The initial and final states: Theoretical implications and experimental explorations of Richness of the Base. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition* (pp. 321–368). Cambridge, UK: Cambridge University Press.
- Davidson, L. (2006a). Phonotactics and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science*, 30, 837–862. DOI: 10.1207/s15516709cog0000_73
- Davidson, L. (2006b). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34, 104–137. DOI: 10.1016/j.wocn.2005.03.004
- Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, 38, 272–288. DOI: 10.1016/j.wocn.2010.01.001
- Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 270–282. DOI: 10.1037/a0020988
- Davidson, L., & Shaw, J.A. (2009). A closer look at perceptual epenthesis in cross-language perception. *Journal of the Acoustical Society of America*, 125, 2752–2753. DOI: 10.1121/1.4784609
- Davidson, L., & Shaw, J.A. (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics*, 40, 234–248. DOI: 10.1016/j.wocn.2011.11.005
- de Lacy, P. (2007). The interaction of tone, sonority, and prosodic structure. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 281–307). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511486371.013

- de Lacy, P. (2006). *Markedness: Reduction and preservation in phonology*. Cambridge, UK: University Press. DOI: 10.1017/CBO9780511486388
- de Saussure, F. (1916). *Cours de linguistique générale*. Lausanne and Paris: Payot. Translated in 1959 as *Course in general linguistics*. New York: Philosophical Library.
- Dell, F., & Elmedlaoui, M. (1985). Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics*, 7, 105–130.
DOI: 10.1515/jall.1985.7.2.105
- Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1355–1367.
DOI: 10.1037/0278-7393.26.6.1355
- Duanmu, S. (2000). *The phonology of Standard Chinese*. Oxford: Oxford University Press.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human perception and performance*, 25, 1568–1578. DOI: 10.1037/0096-1523.25.6.1568
- Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, 16, 491–505.
DOI: 10.1080/01690960143000191
- Eckman, F. R., & Iverson, G. K. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research*, 9, 234–252.
DOI: 10.1177/026765839300900302
- Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. In A. Albright & T. Cho (Eds.), *UCLA Working Papers in Linguistics 7, Papers in Phonology 5* (pp. 71–116). Department of Linguistics of the University of California at Los Angeles.
- Fowler, C., Treiman, R., & Gross, J. (1993). The structure of English syllables and polysyllables. *Journal of Memory and Language*, 32, 115–140. DOI: 10.1006/jmla.1993.1007
- Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics*, 54, 287–295.
DOI: 10.3758/BF03205263
- Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, 42, 481–496. DOI: 10.1006/jmla.1999.2692
- Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory*, 22, 179–228.
DOI: 10.1023/B:NALA.0000005557.78535.3c
- Galves, A., Garcia, J., Duarte, D., & Galves, C. (2002). Sonority as a basis for rhythmic class discrimination. In B. Bel & I. Marlien (Eds.) *Proceedings of the speech prosody 2002 conference* (pp. 323–326). Aix-en-Provence.
- Giegerich, H. J. (1992). *English phonology: An introduction*. New York: Cambridge University Press. DOI: 10.1017/CBO97811391666126
- Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language and Hearing Research*, 42, 708–726. DOI: 10.1044/jslhr.4203.708
- Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition* (pp. 73–108). Cambridge, UK: Cambridge University Press.
- Gordon, M. (2007). Functionalism in phonology. In Paul de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 61–78). Cambridge, UK: Cambridge University Press.

- Gouskova, M. (2001). Falling sonority onsets, loanwords, and syllable contact. In M. Andronis, C. Ball, H. Elston, & S. Neuvel (Eds.), *CLS 37: The Main Session. Papers from the 37th Meeting of the Chicago Linguistic Society*. Chicago: CLS.
- Gouskova, M. (2004). Relational hierarchies in optimality theory: The case of syllable contact. *Phonology*, 21, 201–250. DOI: 10.1017/S095267570400020X
- Greenberg, J.H. (1978). Some generalizations concerning initial and final consonant clusters. In J.H. Greenberg, C.A. Ferguson, & E.A. Moravcsik (Eds.), *Universals in Human Language: 2* (pp. 243–279). Stanford, CA: Stanford University Press.
- Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human perception and performance*, 24, 592–608. DOI: 10.1037/0096-1523.24.2.592
- Hankamer, J., & Aissen, J. (1974). The sonority hierarchy. In A. Bruck, R.A. Fox, & M.W. LaGaly (Eds.), *Papers from the parasession on natural phonology* (pp. 131–145). Chicago: Chicago Linguistic Society.
- Hay, J., Pierrehumbert, J.B., & Beckman, M.E. (2003). Speech perception, well-formedness and the statistics of the lexicon. In J. Local, R. Ogden, & R. Temple (Eds.) *Phonetic interpretation: Papers in laboratory phonology VI* (pp. 58–74). Cambridge, UK: Cambridge University Press.
- Hay, J.A., Pelucchi, B., Graf Estes, K., & Saffran, J.R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, 63, 93–106. DOI: 10.1016/j.cogpsych.2011.06.002
- Hayes, B. (2004). Phonetically-driven phonology: The role of optimality theory and inductive grounding. In J.J. McCarthy (Ed.), *Optimality theory in phonology* (pp. 290–309). Malden: Blackwell Publishing. DOI: 10.1002/9780470756171.ch15
- Hayes, B. (2011). Interpreting sonority projection experiments: The role of phonotactic modeling. *Proceedings of 17th International Congress of Phonetic Sciences*. Hong Kong.
- Hayes, B., & Steriade, D. (2004). Introduction: The phonetic bases of phonological markedness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 1–33). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511486401.001
- Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39, 379–440. DOI: 10.1162/ling.2008.39.3.379
- Hooper, J.B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.
- Jusczyk, P.W., Friederici, A.D., Wessels, J.M.I., Svenkerud, V.Y., & Jusczyk, A.M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402–420. DOI: 10.1006/jmla.1993.1022
- Jusczyk, P.W., Smolensky, P., & Alallo, T. (2002). How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition*, 10, 31–73. DOI: 10.1207/S15327817LA1001_3
- Kabak, B., & Idsardi, W.J. (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech*, 50, 23–52. DOI: 10.1177/00238309070500010201
- Kawasaki-Fukumori, H. (1992). An acoustical basis for universal phonotactic constraints. *Language and Speech*, 35, 73–86.
- Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10, 421–441.

- Kreitman, R. (2006). Cluster Buster: A typology of onset clusters. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 42, 163–179.
- Ladefoged, P. (2001). *A course in phonetics*. Boston: Heinle & Heinle
- Lee, Y. (1994). Onset analysis of Korean on-glides. In Y. K. Kim-Renaud (Ed.), *Theoretical issues in Korean linguistics* (pp. 133–156). Stanford: CSLI.
- Locke, J.L. (2000). Language: Movement patterns in spoken language. *Science*, 288, 449–451. DOI: 10.1126/science.288.5465.449
- MacNeilage, P.F., & Davis, B.L. (2000). On the origin of internal structure of word forms. *Science*, 288, 527–531. DOI: 10.1126/science.288.5465.527
- Massaro, D. W., & Cohen, M. M. (1983). Phonological constraints in speech perception. *Perception & Psychophysics*, 34, 338–348. DOI: 10.3758/BF03203046
- Morelli, F. (1999). *The phonotactics and phonology of obstruent clusters in Optimality Theory*. Doctoral dissertation, University of Maryland, College Park.
- McClelland, J.L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences*, 3, 166–168. DOI: 10.1016/S1364-6613(99)01320-0
- Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, 84, 55–71. DOI: 10.1016/S0010-0277(02)00014-8
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, 25, 83–127. DOI: 10.1017/S0952675708001413
- Moreton, E., Feng, G., & Smith, J.L. (2005). Syllabification, sonority, and perception: new evidence from a language game. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 41, 341–355.
- Nevins, A., & Vaux, B. (2003). Metalinguistic, shmetalinguistic: The phonology of shm-reduplication. *Proceedings of the 39th Chicago Linguistics Society* (pp. 702–721).
- Ohala, D. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders*, 32, 397–422. DOI: 10.1016/S0021-9924(99)00018-0
- Ohala, J.J. (1990). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Chicago Linguistic Society (Ed.) *Papers from the regional meetings* (pp. 319–338). Chicago: CLS.
- Ohala, J. J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. Hakon Jahr (Eds.), *Language and its ecology: Essays in honor of Einar Haugen. Trends in linguistic studies and monographs* (pp. 343–365). Berlin: Mouton de Gruyter.
- Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, 83, B13–B23. DOI: 10.1016/S0010-0277(01)00165-2
- Oudeyer, P-Y. (2005). The self-organization of speech sounds. *Journal of Theoretical Biology*, 233, 435–449. DOI: 10.1016/j.jtbi.2004.10.025
- Padden, C., & Perlmutter, D. (1987). American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory*, 5, 335–375. DOI: 10.1007/BF00134553
- Parker, S. (2002). *Quantifying the sonority hierarchy*. Doctoral dissertation, University of Massachusetts.
- Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, 36, 55–90. DOI: 10.1016/j.wocn.2007.09.003
- Parker, S. (2012). *The sonority controversy*. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110261523

- Pater, J., & Barlow, J. (2003). Constraint conflict in cluster reduction. *Journal of Child Language*, 30, 487–526. DOI: 10.1017/S0305000903005658
- Pelucchi, B., Hay, J., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80, 674–685. DOI: 10.1111/j.1467-8624.2009.01290.x
- Peperkamp, S. (2007). Do we have innate knowledge about phonological markedness? Comments on Berent, Steriade, Lennertz, and Vaknin. *Cognition*, 104, 631–637. DOI: 10.1016/j.cognition.2006.12.009
- Perlmutter, D. (1992). Sonority and syllable structure in American Sign Language. *Linguistic Inquiry*, 23, 407–422.
- Pertz, D. L., & Bever, T. G. (1975). Sensitivity to phonological universals in children and adolescents. *Language*, 51, 149–162. DOI: 10.2307/413156
- Pinker, S., & Birdsong, D. (1979). Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior*, 18, 497–508. DOI: 10.1016/S0022-5371(79)90273-1
- Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics*, 60, 941–951. DOI: 10.3758/BF03211930
- Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishers.
- Redford, M. A. (2008). Production constraints on learning novel onset phonotactics. *Cognition*, 107, 785–816. DOI: 10.1016/j.cognition.2007.11.014
- Ren, J., Gao, L., & Morgan, J. L. (2010). Mandarin speakers' knowledge of the sonority sequencing principle. *20th Colloquium on Generative Grammar*. Barcelona: Universitat Pompeu Fabra.
- Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language*, 64, 83–121. DOI: 10.1006/brln.1998.1958
- Rose, S. (2000). Epenthesis positioning and syllable contact in Chaha. *Phonology*, 17, 397–425. DOI: 10.1017/S0952675701003931
- Sandler, W. (1993). A sonority cycle in American sign language. *Phonology*, 10, 242–279. DOI: 10.1017/S0952675700000051
- Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9781139163910
- Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12, 110–114. DOI: 10.1111/1467-8721.01243
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, 274, 1926–1928. DOI: 10.1126/science.274.5294.1926
- Scholes, R. (1966). *Phonotactic grammaticality*. The Hague: Mouton.
- Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations, part II* (pp. 337–383). Dordrecht: Foris.
- Selkirk, E. O. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oerhle (Eds.), *Language sound structure* (pp. 107–136). Cambridge, MA: MIT press.
- Sibley, D. E., Kello, C. T., Plaut, D. C., & Elman, J. (2008). Large-scale modeling of wordform learning and representation. *Cognitive Science*, 32, 741–754. DOI: 10.1080/03640210802066964
- Smolensky, P. (2006). Optimality in phonology II: Harmonic completeness, local constraint conjunction and feature domain markedness. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar* (pp. 27–160). Vol. 2. Cambridge: The MIT Press.
- Smolensky, P., & Legendre, G. (Eds.) 2006. *The harmonic mind: From neural computation to optimality-theoretic grammar*. Vol. 2. Cambridge: The MIT Press.

- Stenneken, P., Bastiaanse, R., Huber, W., & Jacobs, A. M. (2005). Syllable structure and sonority in language inventory and aphasic neologisms. *Brain and Language*, 95, 280–292. DOI: 10.1016/j.bandl.2005.01.013
- Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. PhD dissertation, MIT.
- Steriade, D. (1988). Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology*, 5, 73–155. DOI: 10.1017/S0952675700002190
- Steriade, D. (2001). *The phonology of perceptibility effects: The P-map and its consequences for constraint organization*. Manuscript, University of California, Los Angeles.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3–45.
- Stoel-Gammon, C. (1985). Phonetic Inventories, 15–24 Months: A longitudinal study. *Journal of Speech and Hearing Research* 28, 505–512. DOI: 10.1044/jshr.2804.505
- Treiman, R. (1984). On the status of final consonant clusters in English syllables. *Journal of Verbal Learning and Verbal Behavior*, 23, 343–356. DOI: 10.1016/S0022-5371(84)90237-8
- Treiman, R., Bowey, J., & Bourassa, D. (2002). Segmentation of spoken words into syllables by English-speaking children as compared to adults. *Journal of Experimental Child Psychology*, 82, 213–238. DOI: 10.1016/S0022-0965(02)00134-0
- Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology*, 33, 771–780. DOI: 10.1037/0012-1649.33.5.771
- Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, 27, 87–104. DOI: 10.1016/0749-596X(88)90050-2
- Treiman, R., Gross, J., & Cwikiel-Glavin, A. (1992). The syllabification of /s/ clusters in English. *Journal of Phonetics*, 20, 383–402.
- Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29, 66–85. DOI: 10.1016/0749-596X(90)90010-W
- Vennemann, T. (1972). On the theory of syllabic phonology. *Linguistische Berichte*, 18, 1–18.
- Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, 36, 481–487. DOI: 10.3758/BF03195594
- Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30, 945–982. DOI: 10.1207/s15516709cog0000_89
- Wilson, C., & Davidson, L. (2013). Bayesian analysis of non-native cluster production. *Proceedings of NELS 40*. Cambridge, MA: MIT.
- Wright, R. (2004). A review of perceptual cues and cue robustness. In D. Steriade, R. Kirchner, & B. Hayes (Eds.), *Phonetically based phonology* (pp. 34–57). Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511486401.002
- Yavas, M., & Gogate, L. J. (1999). Phoneme awareness in children: A function of sonority. *Journal of Psycholinguistic Research*, 28, 245–259. DOI: 10.1023/A:1023254114696
- Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 161–194). Cambridge, UK: Cambridge University Press.
- Zhao, X., & Berent, I. (2013a). Universal constraints on the sound structure of language: Beyond articulation. Poster presented at the *RISE conference*. Northeastern University, Boston.
- Zhao, X., & Berent, I. (submitted). Speakers' knowledge of grammatical universals: Evidence from Mandarin Chinese.
- Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language*, 83, 277–316. DOI: 10.1353/lan.2007.0105

Appendix

The experimental and filler monosyllabic items from Experiments 1–3.

	Filler	Small rise	Plateau	Fall
Fricatives				
	flep	fnik	fsik	msik
	fruk	fnuk	fsuk	msuk
	flet	fnɛt	fset	mset
	frop	fnɔt	fsot	msot
	frip	fɹmik	fɹik	lɹik
	fruk	fɹmuk	fɹuk	lɹuk
	fɹɛp	fɹmɛk	fɹɛk	rɹɛk
	fɹok	fɹmɔk	fɹɔk	rɹɔk
Stops				
	pɹik	pɹnik	pɹtik	mɹtik
	pɹum	pɹnuk	pɹtuk	mɹtuk
	pɹɛf	pɹnɛf	pɹtɛf	mɹtɛf
	pɹɔf	pɹnɔf	pɹtɔf	mɹtɔf
	tɹom	tɹmik	tɹpik	lɹpik
	tɹum	tɹmuk	tɹpuk	lɹpuk
	tɹɛf	tɹmɛk	tɹpɛk	rɹpɛk
	tɹɔk	tɹmɔk	tɹpɔk	rɹpɔk

Corresponding Addresses

Iris Berent
 Department of Psychology
 Northeastern University
 125 Nightingale Hall
 360 Huntington Ave.
 Boston MA 02115
 USA
 i.berent@neu.edu

Tracy Lennertz
 Department of Psychology
 Northeastern University
 125 Nightingale Hall
 360 Huntington Ave.
 Boston MA 02115
 USA
 lennertz.t@gmail.com