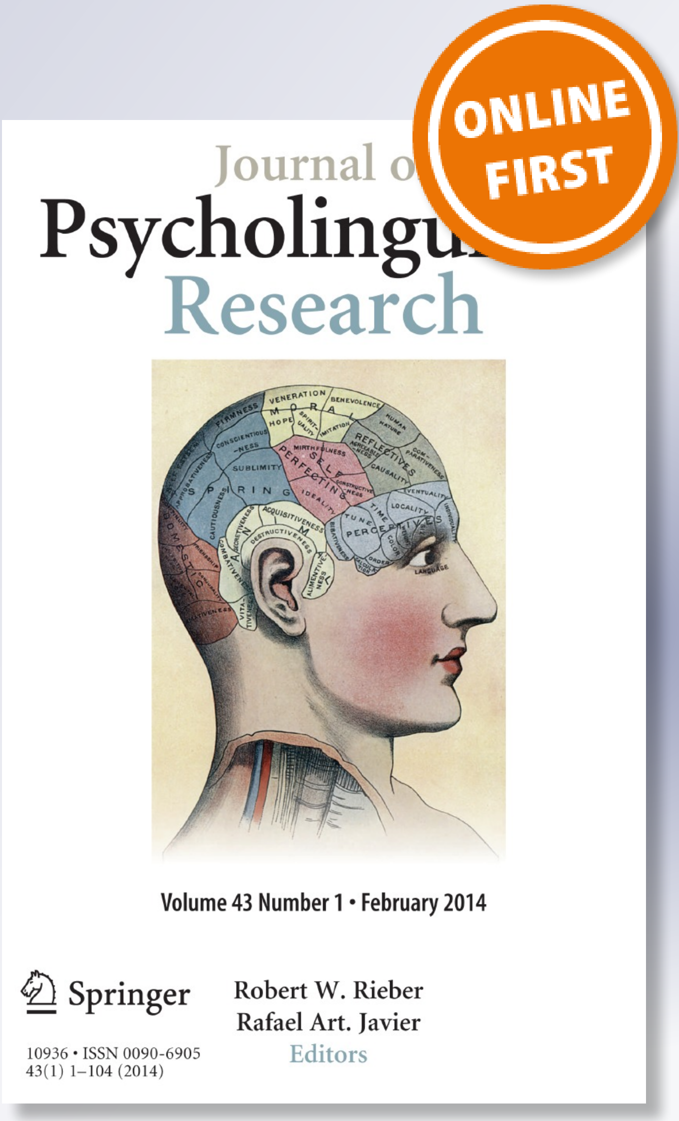
Sensitivity to Phonological Universals: The Case of Stops and Fricatives

Katalin Tamási & Iris Berent

Journal of Psycholinguistic Research

ISSN 0090-6905

J Psycholinguist Res
DOI 10.1007/s10936-014-9289-3



Springer

J Psycholinguist Res
DOI 10.1007/s10936-014-9289-3

Sensitivity to Phonological Universals: The Case of Stops and Fricatives

Katalin Tamási · Iris Berent

Abstract Linguistic evidence suggests that syllables like *bdam* (with stop–stop clusters) are less preferred than *bzam* (with stop–fricative combinations). Here, we demonstrate that English speakers manifest similar preferences despite no direct experience with either structure. Experiment 1 elicited syllable count for auditory materials (e.g., does *bzam* have one syllable or two?); Experiment 2 examined the AX discrimination of auditory stimuli (e.g., is *bzam* = *bezam*?); whereas Experiment 3 repeated this task using printed materials. Results showed that syllables that are dispreferred across languages (e.g., *bdam*) were prone to misidentification relative to preferred syllables (e.g., *bzam*). The emergence of this pattern irrespective of stimulus modality—for auditory and printed materials—suggests that misidentification does not solely stem from a phonetic failure. Further, the effect remained significant after controlling for various statistical properties of the materials. These results suggest that speakers possess broad linguistic preferences that extend to syllables they have never encountered before.

Keywords Phonological-universals · Phonology · Reading · Sonority · Optimality-theory

Introduction

Natural languages are known to exhibit systematic regularities in the distribution of syllable structures. Across languages, certain syllables (e.g., *lbif*) are less frequent than others (e.g., *bnif*; Berent et al. 2007; Greenberg 1978). Past research has demonstrated that these regularities converge with the behavior of individual speakers, as structures that are underrepresented across languages also tend to be dispreferred by individual speakers (Berent et al. 2008; Broselow and Finer 1991; Fleischhacker 2005; Greenberg and Jenkins 1964; Pertz and Bever 1975). But whether this convergence is robust, and whether

K. Tamási · I. Berent (✉)
Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Ave, Boston, MA 02115, USA
e-mail: i.berent@neu.edu

Published online: 24 February 2014

Table 1 Sonority scale of speech sounds

Sound category	Class	Example	Sonority level
Sonorants	Vowels	*a*, *i*	6
	Glides	*y*, *w*	5
	Liquids	*l*, *r*	4
	Nasals	*m*, *n*	3
Obstruents	Fricatives	*v*, *z*	2
	Stops	*b*, *d*	1

it is due to universal grammatical constraints or non-grammatical sources (e.g., sensorimotor pressures and statistical knowledge; Blevins 2006; Bybee and McClelland 2005; Byrd 1992; Davidson 2010, 2011a,b, 2012; Dupoux et al. 2011; Redford 2008; Saffran et al. 1996; Vitevich and Luce 2005; Wright 2004) remains open empirical questions.

In what follows, we further address these issues by investigating a new case of a putatively universal restriction on syllable structure. We first briefly review a grammatical account for this phenomenon and introduce our case study, our manipulations and results. The General Discussion considers competing explanations for the findings.

Sonority Restrictions on Syllable Structure

Our investigation specifically concerns the restrictions on onset clusters—the string of consonants that occur at the beginning of the syllable (e.g., *bl* in *black*). As noted above, across languages, certain onset clusters (e.g., *bla)* are preferred to others (e.g., *lba)*. Linguistic analyses capture these facts by sonority restrictions (Clements 1990; Parker 2002, 2008; Selkirk 1984; Steriade 1982).

Sonority is an abstract phonological feature that correlates with intensity (Ladefoged 2001). Each speech sound can be categorized in terms of its sonority level (see Table 1):[1] most sonorous sounds are vowels, followed by glides (e.g., *y*, *w*), liquids (e.g., *l*, *r*) and nasals (e.g., *m*, *n*), which together form the class of sonorants. Next on the scale are obstruents—a group that comprises fricatives (e.g., *v*, *z*) and, finally, stops (e.g., *b*, *d*)—the least sonorous on the scale.

Using this scale, one can further compute the sonority distance of an onset cluster by subtracting the sonority level of the first consonant from that of the second ($\Delta s = S_2 - S_1$). In the case of *bl*, the sonority distance yields a large positive number ($\Delta s = 4 - 1 = 3$). Following the same principle, onsets such as *bn* manifest a smaller rise in sonority ($\Delta s = 2$), onsets like *bd* exhibit a sonority plateau ($\Delta s = 0$), whereas *lb*-type onsets fall in sonority ($\Delta s = -3$).

While all languages constrain the sonority profile of the syllable, distinct languages differ on the range of sonority distances that they allow. English requires its onsets to exhibit a large sonority rise—it allows onsets like *bl* ($\Delta s = 3$), but not *bn, bd* or *lb* ($\Delta s = 2$, $\Delta s = 0$, $\Delta s = -3$, respectively). Other languages like Albanian or Russian tolerate even negative sonority distances (e.g., *lb*, $\Delta s = -3$ Gouskova 2001; Klippenstein 2008). But this cross-linguistic variation is nonetheless systematic: languages that tolerate onsets with smaller sonority distances tend to allow larger distances (e.g., *lb* $\Rightarrow$ *bd*), whereas

[1] The linguistic literature has proposed various sonority scales that differ in detail, ranging from five (Clements 1990) to seventeen levels (Parker 2008). For the sake of simplicity, we follow Selkirk (1984) and Parker (2002) in distinguishing the sonority levels of stops and fricatives, but in other respects, we use the rudimentary sonority scale proposed by Clements (1990). Our analyses disregard complex obstruent affricates (containing a stop and a fricative, e.g., the first sound in *Joe*) and treat sonority as an ordinal scale.

languages that exhibit large sonority distances do not necessarily allow smaller ones (data from Greenberg 1978, reanalyzed by Berent et al. 2007). These observations suggest a cross-linguistic hierarchy of onset clusters: large sonority distances are preferred to smaller ones. Specifically, *bl* ≻ *bn* ≻ *bd* ≻ *lb* (where ≻ indicates preference, Berent et al. 2007).

Optimality Theory (Prince and Smolensky 1993/2004) attributes this hierarchy to universal grammatical constraints that favor large sonority distances over smaller ones (Smolensky 2006).[2] By hypothesis, these constraints are present in the grammar of every speaker, irrespective of whether the relevant clusters are present in their language or absent. The existing experimental findings are consistent with this prediction.

Past Experimental Evidence for Sonority Restrictions

Past research has shown that people generally favor onsets with large sonority distances (e.g., *bla*); such onsets are acquired earlier in both first- (Barlow 2001, 2005; Gierut 1999; Ohala 1999; Yavas and Gogate 1999) and second-language (Eckman and Iverson 1993) and they are more likely to be preserved in aphasia (Christman 1992; Romani and Calabrese 1998; Stenneken et al. 2005). Furthermore, people systematically extend the sonority hierarchy even to onsets that they have never heard before (Berent et al. 2007, 2008, 2009, 2010, 2011a,b,c; Berent and Lennertz 2010; Lennertz and Berent 2013; Zhao and Berent 2013).

The critical evidence comes from a phenomenon of perceptual illusions. Previous research demonstrated that clusters that are unattested in one's language tend to be misidentified (Dupoux et al. 1999; Massaro and Cohen 1983; Moreton 2002; Pitt 1998). For example, English speakers tend to misidentify the unattested *dla* as *dela*; (Pitt 1998). Berent et al. (2007) hypothesized that misidentification has a grammatical origin: ill-formed onsets undergo repair in order to abide by universal grammatical restrictions—the worse-formed the onset, the more likely the repair. Such restrictions might include universal grammatical constraints on sonority.

To test this possibility, Berent et al. (2007) examined the identification of various types of onset clusters, ranging from small rises in sonority (e.g., *bnif*) to sonority plateaus (e.g., *bdif*) and falls (e.g., *lbif*). Results showed that, as sonority distance decreased, people were more likely to misidentify the monosyllable (e.g., *lbif*) with its disyllabic counterpart (e.g., *lebif*). Remarkably, the sensitivity to onset structure obtained despite the fact that none of these clusters were attested in participants' language (English).

Additional results suggested that these perceptual illusions are not solely due to the similarity of these onsets to attested English words, as the findings replicate with speakers of Korean and Chinese—languages that ban onset clusters altogether (Berent et al. 2008; Zhao and Berent 2013). It is also unlikely that misidentification is due to the failure to extract the phonetic form of auditory onsets (Dupoux et al. 2011). First, English participants are demonstrably able to correctly encode auditory ill-formed onsets (e.g., *mdif*) under conditions that promote attention to phonetic form (Berent et al. 2007, 2011c). Moreover, the misidentification of ill-formed onsets obtains even with *printed* materials (Berent and Lennertz 2010; Berent et al. 2009; Lennertz and Berent 2013). These results suggest that misidentification reflects neither phonetic failure nor lexical unfamiliarity. Instead, misidentification might result from active grammatical repair, triggered by the grammatical ill-formedness of the onset.

[2] The restrictions on onset structure can acquire multiple forms—some directly appeal to sonority, whereas others do not (Smolensky 2006). We remain agnostic as to the exact representation of sonority restrictions in the language system—whether sonority is represented as a scalar phonological feature (c.f., Clements 1990) or whether it results from other constraints on feature conjunction (Smolensky 2006). Our question here is whether sonority can be used descriptively, to capture the well-formedness of the onset.

Sonority Levels of Fricatives and Stops

Although there is much evidence to suggest that people are sensitive to sonority distance, most of the existing evidence comes from onsets with relatively large sonority distances, such as the clines between obstruents and sonorants (e.g., *bnif*). Linguistic research, however, suggests that the class of obstruents comprises of two distinct sub-categories—stops and fricatives. Moreover, fricatives, on this account are more sonorous than stops. If people possess universal sonority restrictions, then they might extend them even to the slight sonority clines in stop–fricative combinations. Because stops are less sonorous than fricatives, stop–fricative onsets (e.g., *bz*) should exhibit a slight rise in sonority, hence, they should be better formed than stop–stop and fricative–fricative sequences (i.e., plateaus, e.g., *bd and zv*, respectively), which in turn, should be favored to fricative–stop combinations (i.e., sonority falls, e.g., *zb*).

Linguistic analyses are consistent with this prediction. Consider, for example, the syllabification of words in the Imdlawn Tashlhiyt dialect of Berber (a language spoken in Northern Africa). It is well known that syllables require their nuclei to exhibit a sonority peak. While most languages limit the nucleus to a vowel (e.g., *b**a**g*), Imdlawn Tashlhiyt allows obstruents, such as *t**Z**.di* ('put together') or *ra.t**K**.ti* ('she will remember'; capitalizations denote the nucleus; periods denote syllable boundaries). Crucially, fricative nuclei are preferred to stops. Accordingly, the word *tftkt*, 'she suffered a sprain' is syllabified as *t**F**.tkt* (with a fricative nucleus) rather than *tf**T**.kt* (with a stop nucleus) (Dell and Elmedlaoui 1985). Further evidence for the stop/fricative sonority distinction comes from cluster reductions in first language acquisition (Gnanadesikan 2004; Ohala 1999) and productive phonological processes in languages like Ancient Greek (Steriade 1982).

English, however, does not systematically distinguish between the sonority levels of stops and fricatives. Most onsets allow stops and fricatives to combine with the same set of segments (e.g., *ply* vs. *fly*, *brand* vs. *friend*, [*bj*]*uty* vs. [*fj*]*uel*). The only counter-example concerns the segment *s*—the only English obstruent to combine with another obstruent (e.g., *stake, sport*). This segment is known to systematically violate sonority restrictions in many languages (Steriade 1982), and it will not be discussed further or included our experimental manipulations.[3] Putting aside the case of word-initial *s*, we ask whether English speakers are nevertheless sensitive to the minute sonority clines between stops and fricatives (e.g., *bza* vs. *zba*).

Previous research (Lennertz and Berent 2013) has examined the sonority of stops and fricatives indirectly, by comparing the size of the sonority clines in fricative–nasal and stop–nasal onsets (e.g., *pn* vs. *fn*). To control for the distinct phonetic demands of processing stops and fricatives, each such sequence was compared to a sonority plateau baseline, matched for the initial consonant: *pt* was compared to *fs; pn* was compared to *fn.* If fricatives are more sonorous than stops, then the sonority cline in fricative–nasal onsets (e.g., *fn*) should be smaller (i.e., worse-formed) than stop–nasal sequences (e.g., *pn*), hence, *fn*-onsets should be more likely to elicit misidentification. Results from several experiments were consistent with this prediction. Other studies, however, found no sensitivity to the sonority profile of stop–fricative onsets (Davidson 2011a), but those studies did not control for the phonetic properties of the initial segment (e.g., the presence of a release burst in stops), so it is conceivable that these properties could have masked the effect of sonority. Accordingly, the existing results do not establish whether English participants distinguish the sonority profile of stops and fricatives.

[3] Note that such counterexamples would incorrectly suggest that fricatives are *less* sonorous than stops, as segments allowed in the first onset position are typically more sonorous than their C2 counterparts.

Table 2 The design of Experiments 1 and 2 (Auditory stimuli)

Obstruent type	Sonority distance	
	Better-formed	Worse-formed
Stop-initial	*bzam*	*bdam*
Fricative-initial	*vzam*	*vdam*

The Present Study

The present research further investigates whether English speakers are sensitive to the small sonority clines between stops and fricatives. The key difference between our approach and past research (Lennertz and Berent 2013) concerns the design of the material. While past research gauged the sonority distance of such clusters indirectly, by comparing stop–nasal and fricative–nasal sonority rises to their respective plateau baselines (e.g., *pn* vs. *pt* and *fn* vs. *fs*), our studies combine stops and fricatives in the same onset to create obstruent clusters (e.g., *bz, bd*). This allows us to directly examine whether participants are sensitive to the structure of obstruent–obstruent onsets.

The experimental stimuli featured three types of onsets, defined by their sonority distance of the onset. In the first type of items, stops are followed by fricatives (e.g., *bz*) to yield a slight sonority rise. The second type exhibits a sonority plateau comprising either two stops (e.g., *bd*) or two fricatives (e.g., *vz*). Finally, in the third type, fricatives precede stops to form a sonority fall (e.g., *vd*).

We ask whether English speakers can differentiate between such minute sonority distances (i.e., between small rises and plateaus or between plateaus and small falls). As in previous research, we infer people's sensitivity to onset structure from their tendency to misidentify ill-formed onsets (Berent et al. 2007, 2008, 2009, 2010, 2011a,b,c; Berent and Lennertz 2010; Lennertz and Berent 2013; Zhao and Berent 2013). Our experiments test two pairwise contrasts in sonority distances (see Table 2). Each pair is matched for the initial consonant, and their sonority distance is manipulated. The first contrast pits sonority plateaus (e.g., *bdam*) against sonority rises (e.g., *bzam*). The second contrast pits sonority falls (e.g., *vdam)* and plateaus (e.g., *vzam*). If small sonority distances are ill-formed, then the smaller sonority distance in the first pair member should render it worse-formed, hence, more vulnerable to grammatical repair than its counterpart: *bdam* should be more likely to be misidentified than *bzam*; *vdam* should be more prone to repair than *vzam*. Experiments 1–2 test this prediction using auditory stimuli. To determine whether the misidentification of ill-formed onsets is due to their acoustic properties, Experiment 3 uses printed materials.

Experiment 1

The first study investigates whether English speakers are sensitive to the small sonority distance in stop–fricative onsets using a syllable-count task. In each trial, participants were presented with a single auditory stimulus—either a monosyllable with a complex onset (e.g., *bzam*) or its disyllabic counterpart (e.g., *bezam*). Their task was to judge whether the stimulus they heard had one syllable or two. The critical manipulation concerns the sonority distance of the onsets. Onsets were either better-formed (small rises: *bz*) or worse-formed (plateaus: *bd*). To partly control for the phonetic properties of the initial obstruent, better- and worse-formed onsets were matched for the initial consonant, and their type was manipulated—either a stop or a fricative. Accordingly, our experiment examined the effect of sonority along two comparisons. One comparison pitted small rises against plateaus (e.g., *bzam* vs. *bdam*; another comparison contrasted sonority plateaus and falls (e.g., *vzam* vs. *vdam*). We expect worse-formed monosyllabic items to be more likely to undergo repair, hence, they should be harder to identify as monosyllables than better-formed items. Specifically, given stop-initial onsets, sonority plateaus should be worse-formed, hence, harder to identify than rises. Likewise, repair should be more likely for the fricative-initial sonority falls (e.g., *vdam*) than the fricative-initial plateaus (e.g., *vzam*), hence, falls should be more prone to misidentification than plateaus.

Method

Participants

Sixteen native English speakers, undergraduate students at Northeastern University, participated in this experiment in partial fulfillment of a course requirement.

Materials

The experimental materials consisted of 48 monosyllabic (e.g., *bdam*) and 48 disyllabic (e.g., *bedam*) items (for the list of the monosyllabic items see Appendix 1). Monosyllables were arranged in quartets, generated by crossing two variables of interest (1) the type of the initial consonant: stop versus fricative; and (2) the sonority distance of the onset—either better-formed with a larger sonority distance or worse-formed with a smaller sonority distance (for the factorial design, see Table 2). In stop-initial items, better-formed onsets manifested a sonority rise, whereas worse-formed onsets had a sonority plateau (e.g., *bzam* vs. *bdam*). In fricative-initial-items, the better-formed onsets had a sonority plateau, whereas worse-formed onsets had a sonority fall (e.g., *vzam* vs. *vdam*). Quartet members were matched for their rhyme, and they contrasted on their onset structure (e.g., *bzam, bdam, vzam, vdam*).

To assure that the effect of these two variables of interest was not tainted by other properties, unrelated to sonority, we also controlled our materials for several linguistic aspects. First, onset consonants were always heterorganic (i.e., they had different places of articulation)—this control was instituted because homorganic consonants are typically banned across languages (McCarthy 1986). Second, since labial–coronal ordering is cross-linguistically favored to the coronal–labial one (Byrd 1992), all onsets began with a labial consonant (i.e., *b* or *v*), followed by a coronal consonant (*d* or *z*). Third, onset consonants were matched for voicing, since across languages, voiceless consonants are less sonorous than their voiced counterparts (Parker 2002; Steriade 1982). Finally, to minimize the feature similarity between onset and coda consonants, we selected onset (*b*, *v*, *d*, *z*) and coda (*m*, *n*, *g*) consonants from distinct non-overlapping sets.

Disyllables differed from monosyllables in one crucial respect: they contained a schwa (i.e., epenthesis) between the two initial consonants (e.g., *bəzam*). The monosyllabic items and their epenthetically related disyllabic counterparts were selected to closely match each other in terms of pitch contour and overall voice quality by inspecting their spectrogram using Praat (Boersma and Weenink 2003) and by auditory inspection. To ensure that participants clearly hear the onset, we inserted a 50 ms silence at the beginning of each stimulus item.

The materials were recorded by a female native Russian speaker (Russian allows all these onset types, so those stimuli can be produced naturally). The recording lists paired monosyllables with their disyllabic counterparts, counter-balanced for order (e.g., *bzam-bezam; bezam-bzam*). The speaker was instructed to produce the pair members as similarly to each other as possible, and maintain the same intonation throughout.

In order to familiarize participants with the task, they were first given practice with English words, consisting of monosyllables with onset clusters and similar disyllabic stimuli (e.g., *sport, support, crate, curate, blow, below, drive, derive*).

Stimulus Validation

To ensure that our monosyllabic and disyllabic stimuli were indeed produced as intended, we asked five native Russian speakers to complete the auditory syllable count task (Experiment 1) and the discrimination task (Experiment 2) in a counterbalanced order (data from one additional participant was excluded because he reported difficulties understanding the task, and his overall performance was close to chance level, $M = 54\,\%$).

Participants correctly identified the monosyllabic ($M = 93\,\%$) and disyllabic ($M = 80\,\%$) items with high accuracy (i.e., accurate responses are defined as ones that are consistent with the talker's intention). Given the small sample size, we only ran the analyses using items as a random variable. A 2 OBSTRUENT TYPE (i.e., stop-initial, fricative-initial) × 2 SONORITY DISTANCE (i.e., smaller vs. larger distance) ANOVA on response accuracy to the monosyllabic items yielded no significant effects or interaction (all $F \leq 1.44$, $p \geq .26$). A similar analysis conducted on the disyllabic items revealed no significant effects or interaction (all $F \leq 4.27$, $p \geq .094$). These results demonstrate that our monosyllabic and disyllabic items are perceptible as such, regardless of sonority distance.

Procedure

After providing their informed consent, participants wore headphones and sat in front of the computer.

Each trial began with a screen including the fixation point (*), the trial number and a prompt to press the space-bar to begin the trial. When participants initiated the trial, they were presented with an auditory stimulus. Participants were required to rapidly determine whether a given stimulus contained one or two syllables and to indicate their response by pressing the appropriate key ("1" = one syllable, "2" = two syllables). Prior to the experiment, participants were provided with a short practice session. During practice, participants received accuracy feedback (the words "correct" or "incorrect" flashed on the screen following the trial). Slow responses (>2,500 ms) triggered a warning message from the computer ("too slow"). During the experimental session, only response time feedback was provided. Both practice and experimental trials were presented in a randomized order and the whole task took about 20 min. Participants were run on Experiments 1 and 2 in a counterbalanced order. The experimental procedure was carried out with E-prime software (Schneider et al. 2002).

Results

Responses provided to the monosyllabic and disyllabic items were analyzed separately. In this and all subsequent experiments, correct responses given faster than 200 ms or slower than 2.5 SD of the mean response time were treated as outliers, and they were excluded from the analysis of response time. Outliers comprised 3 % of the trials.

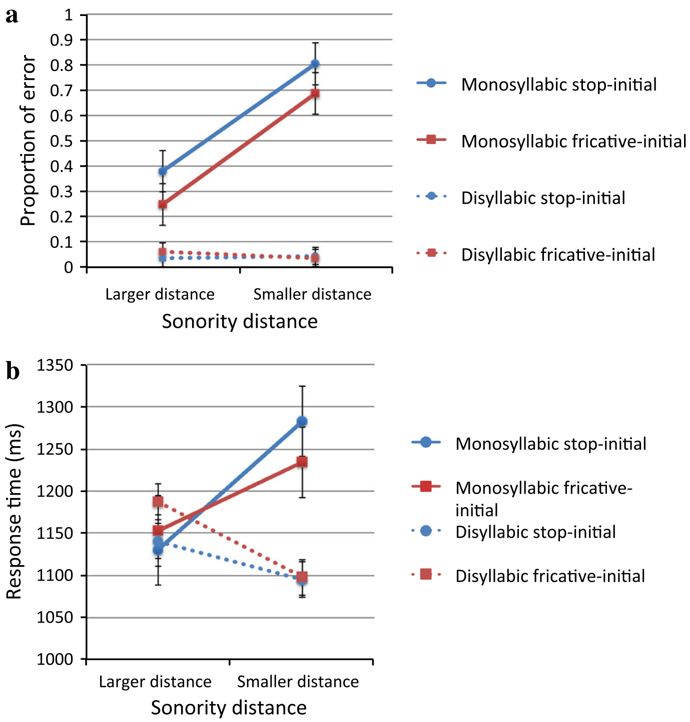


Fig. 1 Mean proportion of error (**a**) and mean response time (ms) (**b**) of the monosyllabic and disyllabic items in Experiment 1 as a function of sonority distance. *Error bars* reflect the confidence intervals constructed for the difference among the means

Responses to Monosyllabic Items

Figure 1 provides the mean response time and errors to monosyllables. An inspection of the means suggested that better-formed items (i.e., those with larger sonority distance) were identified more accurately and more quickly than worse-formed ones (those with smaller distance), and the advantage of better-formed onsets was not further modulated by obstruent type. This observation was confirmed by 2 OBSTRUENT TYPE (i.e., stop-initial vs. fricative-initial items) × 2 SONORITY DISTANCE (smaller vs. larger distance) ANOVAs, using both participants (F1) and items (F2) as random variables.[4] These analyses yielded a significant main effect of sonority distance in response accuracy (F1(1, 15) = 36.455, MSE = .082, $p < .001$), F2(1, 11) = 226.513, MSE = .010, $p < .001$) and response time (F1(1, 10) = 7.245, MSE = 20,802, $p < .02$, F2(1, 9) = 8.059, MSE = 42,729, $p < .01$).

The ANOVAs also yielded a significant effect of obstruent type in response accuracy (F1(1, 15) = 16.265, MSE = .02, $p < .001$, F2(1, 11) = 11.155, MSE = .016, $p < .007$; In response time: F1(1, 10) = .1308, MSE = 14,657, $p < .73$, F2(1, 9) = 4.021, MSE = 13,492, $p < .08$), as fricative-initial items were identified more accurately than stop-initial items.

[4] In the analysis of response time, the exclusion of trials faster than 200 ms and slower than 2.5 SD of mean response time yielded missing cells. By applying list-wise deletion, 4 participants with missing data were excluded from the subject analyses (N = 11) and 2 quartets with missing data were excluded from the item analyses (N = 10).

However, interaction was not significant in either response accuracy or in response time (all $F \leq 1.43$, $p \geq .26$).[5]

Responses to Disyllabic Items

An inspection of Fig. 1 further suggests that participants identified the disyllabic counterparts of worse-formed monosyllables (e.g., *bedam*, counterpart of *bdam*) faster than the counterparts of better-formed monosyllables (e.g., *bezam*, counterpart of *bzam*). The 2 OBSTRUENT TYPE X 2 SONORITY DISTANCE ANOVAs on responses to disyllables indeed yielded a reliable main effect of sonority distance ($F1(1, 15) = 13.347$, $MSE = 5{,}498$, $p < .002$, $F2(1, 11) = 15.82$, $MSE = 2{,}677$, $p < .002$). The obstruent type factor was not significant, nor did it interact with sonority distance (all $F \leq 2.7$, $p \geq .121$). Similar analyses conducted on the proportion of errors yielded no reliable effects (all $F \leq .32$, $p \geq .58$).[5]

Discussion

Experiment 1 examined whether English speakers encode the small sonority clines in obstruent–obstruent onsets consisting of stop–fricative combinations. To this end, we manipulated the sonority distance in pairs of monosyllabic items, either stop- (e.g., *bzam* and *bdam*, respectively) or fricative-initial items (e.g., *vzam* and *vdam*, respectively). Results suggest that speakers are sensitive to the structure of such onsets. Worse-formed monosyllables systematically elicited slower and inaccurate responses relative to better-formed monosyllables, and this effect obtained irrespective of the initial consonant—stop or fricative.

In fact, the sonority distance of the monosyllable even affected responses to their disyllabic counterparts. Although our disyllabic stimuli were all possible English words, we found that the disyllabic counterparts of better-formed monosyllables required additional processing. That is, it took longer to identify *bezam* (counterpart of the better-formed *bzam*) as disyllabic compared to *bedam* (counterpart of the worse-formed *bdam*). This effect does not appear to stem from the phonetic properties of the disyllables, specifically, the duration of their schwa (the element that distinguishes disyllables from the monosyllable). This conclusion is supported by auxiliary step-wise linear regression analyses that examined the unique contribution of schwa duration[6] and sonority distance as two ordered predictors. When forced last into the model, schwa duration did not capture significant unique variance in either response time or accuracy (see Table 3). In contrast, when the order of predictors was flipped, the unique effect of sonority distance (entered last) remained significant in the analysis of response time even after controlling for schwa duration (see Table 3).

While the effect of sonority distance on disyllables is not captured by their own phonetic properties, this finding (replicating past results in English and Korean, c.f., Berent et al. 2008) can be explained by the phonological properties of their monosyllabic counterparts.

[5] All accuracy results were supported by a mixed-effect logit model, with obstruent type and well-formedness as fixed effects (both sum coded) and subject and quartet as a random effects (R Development Core Team 2011). The results confirmed the effect of sonority distance ($\beta = -1.13$, $SE = .09$, $Z = -12.251$, $p < 2^{e} - 16$) and obstruent type ($\beta = -.36$, $SE = 0.1$, $Z = -3.47$, $p < .0005$) and no interaction ($\beta = -.008$, $SE = .09$, $Z = -.09$, $p < .93$). Similar models for disyllabic trials revealed no significant effects or interaction (all $\beta = .22$, $p = .24$).

[6] In stop-initial items, we defined the beginning of the vowel as the zero-crossing before the change in waveform amplitude and formant structure associated with the vowel, thus excluding stop closure and release. In fricative-initial items, we excluded fricative turbulence preceding the vowel. The end of the vowel was taken to be the zero-crossing before the stop closure and release (if the vowel was followed by a stop) and fricative turbulence (if it was followed by a fricative).

Table 3 Step-wise linear regression analyses examining the contribution of the duration of schwa on response accuracy and response time in Experiment 1

Last predictor	Predictors forced in previous steps	R^2_{change}	F_{change}	*df*	$p <$
a. Response accuracy					
a. Sonority distance	Duration, obstruent type	0.014	0.658	1, 44	NS
b. Duration	Sonority distance, obstruent type	0.078	3.754	1, 44	NS
b. Response time					
a. Sonority distance	Duration, obstruent type	0.171	9.36	1, 44	0.021
b. Duration	Sonority distance, obstruent type	0.005	0.248	1, 44	NS

Upon hearing *bezam*, participants must determine whether they have heard a monosyllable or a disyllable (i.e., *bzam* or *bezam*). Because *bzam* is better formed, it competes with the correct response (*bezam*), more effectively than the worse formed *bdam* competes with its counterpart *bedam*.

Nonetheless, some of our results appear to reflect systematic effects unrelated to sonority distance. In particular, fricative-initial monosyllabic items were overall identified more accurately than stop-initial items. The misidentification of stop-initial monosyllables may have a phonetic basis. Indeed, stop-initial items are inherently discontinuous (c.f., Stevens 1989), and our past research (Berent and Lennertz 2010; Lennertz and Berent 2013) has shown that English speakers tend to interpret such discontinuity as evidence for bipartite structure, hence, disyllabicity. This discontinuity could also account for the misidentification of stop-initial monosyllables in the present experiment.

Taken as a whole, the results of Experiment 1 suggest that the sonority distance between stops and fricatives modulated English speakers' behavior: monosyllables with better-formed onsets were identified more accurately and more quickly than worse-formed items, and the structure of the monosyllable even affected responses to their disyllabic counterparts. These findings are consistent with the hypothesis that English speakers are sensitive to the distinction between the sonority levels of fricatives and stops.

Experiment 2

The results of Experiment 1 show that English speakers misidentify worse-formed monosyllabic items (e.g., *bdam*) as disyllables (e.g., *bedam*). The susceptibility of such monosyllables to misidentification is in line with the hypothesis that such items are repaired as disyllables. Experiment 2 directly tests this possibility by asking participants to discriminate those monosyllables from their disyllabic counterparts. To that end, participants were presented with item pairs—either two identical items (e.g., two monosyllables, e.g., *bdam-bdam*, or two disyllables: *bedam-bedam*) or nonidentical items that paired monosyllables with their disyllabic counterparts (e.g., *bdam-bedam*). Participants were asked to determine whether the pair members were identical.

In line with Experiment 1, we predicted that discrimination should be modulated by sonority distance. That is, monosyllables with worse-formed onsets (i.e., those with smaller sonority distance, e.g., *bdam*) should be more prone to grammatical repair, and consequently, they will be more likely to be erroneously judged as identical to their disyllabic counterparts (e.g., to *bedam*) relative to better-formed items (e.g., in *bzam-bezam*).

Method

Participants

Sixteen native English speakers, undergraduate students at Northeastern University, participated in this experiment in partial fulfillment of a course requirement. These participants also took part in Experiment 1 (the experiments were administered in a counterbalanced order). One participant was excluded (only in Experiment 2) because his response to identical monosyllables (both response time and errors) fell more than 2 standard deviations below the group mean.

Materials

The stimuli were identical to the ones in Experiment 1. The stimulus items were arranged in pairs. In half of the trials, the members of the pair were identical tokens, either monosyllabic (e.g., *bdam-bdam*) or disyllabic (*bedam-bedam*). The other half of the trials contained non-identical, epenthetically related stimuli, with their order counterbalanced (*bdam-bedam, bedam-bdam*).

Two lists of stimulus pairs were created such that each list presented each stimulus in the first position exactly once. The two lists were balanced in terms of identity (identical/non-identical stimuli), sonority distance (worse-formed/better-formed), initial consonant (stop/fricative) and presentation order (i.e., monosyllabic items occurred half of the time in the first position in both lists). Each list included 96 experimental trials. Participant assignment was counterbalanced between the two experimental lists.

The structure of the practice session was similar, and it consisted of the same items as in Experiment 1 (a total of 8 trials).

Procedure

After providing their informed consent, participants were seated in front of the computer and they wore headphones.

Participants initiated the trials by pressing the space bar. Their responses triggered the presentation of the first member of the pair. The second stimulus followed with a stimulus onset asynchrony (SOA) of 1,500 ms. Participants were instructed to determine as quickly and accurately as possible whether the two items were identical or not and indicate their responses by pressing one of two keys (“1” key if they judge the two items to be identical, and “2” if they judge them to be non-identical). Slow responses (RT $>$ 2,500 ms) received a computerized warning signal (“too slow”).

Prior to the experimental session, a practice session was administered in order to familiarize the participants with their task. In addition to feedback on response time, participants received accuracy feedback (the words “correct” or “incorrect” flashed on the screen following the trial) during the practice session.

Results and Discussion

Identical (e.g., *bzam-bzam*) and non-identical trials (e.g., *bzam-bezam*) were analyzed separately. Outliers consisted of 3 % of the total correct responses.

Table 4 Mean proportion of error and mean response time (ms) of identical trials in Experiment 2 as a function of sonority distance and obstruent type

	Better-formed	Worse-formed
Mean proportion of error		
Stop-initial trials	.03 (.27)	.02 (.06)
Fricative-initial trials	.11 (.31)	.08 (.12)
Mean response time		
Stop-initial trials	1,082 (146)	1,048 (126)
Fricative-initial trials	1,080 (122)	1,061 (152)

Standard deviations are indicated in parentheses

Identical Trials

The 2 SYLLABLE (i.e., monosyllabic-monosyllabic or disyllabic–disyllabic) × 2 OBSTRUENT TYPE (i.e., stop-initial vs. fricative-initial) × 2 SONORITY DISTANCE (smaller vs. larger distance) ANOVAs on response to identical trials yielded no reliable effects (for the means, see Table 4). Specifically, no effects were significant in the analyses of errors (all $F < 1.93$, $p = .19$).[7] In response time, the only effects to approach significance were the three-way interaction ($F1(1, 15) = 6.06$, $MSE = 4{,}082$, $p < .03$, $F2(1, 11) = 2.31$, $MSE = 5{,}942$, $p < .16$) and the main effect of sonority distance ($F1(1, 14) = 7.65$, $MSE = 2{,}783$, $p < .03$, $F2(1, 11) = 2.45$, $MSE = 6{,}381$, $p < .15$). These effects, however, were not significant by items. For the means, see Table 4.

Non-Identical Trials

An inspection of the means (see Fig. 2) revealed that trials with better-formed items (e.g., *bzam–bezam*) were more accurately classified than those with worse-formed items (e.g., *bdam–bedam*), and this was so regardless of presentation order (e.g., *bdam–bedam* vs. *bedam-bdam*). The 2 PRESENTATION ORDER (i.e., monosyllabic–disyllabic or disyllabic–monosyllabic) × 2 OBSTRUENT TYPE (i.e., stop-initial vs. fricative-initial) × 2 SONORITY DISTANCE (i.e., smaller vs. larger distance) ANOVAs indeed yielded a reliable effect of SONORITY DISTANCE ($F1(1, 15) = 3.78$, $MSE = .06$, $p < .001$, $F2(1, 11) = 34.56$, $MSE = .04$, $p < .001$). No other effects or interactions reached significance (all $F \leq 2.45$, $p \geq .15$). Likewise, there were no reliable effects or interactions in the response time measure (all $F \leq 1.85$, $p \geq .21$).

The susceptibility of monosyllables with small sonority distance to misidentification suggests that such onsets are encoded as disyllables. Finding that misidentification persists even when monosyllables are explicitly compared to their disyllabic counterparts could suggest that the erroneous encoding of such ill-formed onsets is automatic.

Experiment 3

Why are onsets with small sonority distance vulnerable to misidentification? Earlier, we suggested that misidentification reflects an active process of grammatical repair. In this view, monosyllables are actively recoded as disyllables in order to abide by grammatical constraints that ban small sonority distance—the smaller the distance, the more likely the recoding. But on

[7] The mixed effect logit models on response accuracy did not support the trends observed in ANOVAs (all *it* $\beta = .08$, $p = .69$). In the response time measure, no effects or interactions were significant (all $\beta = .11.5$, $p = .25$).

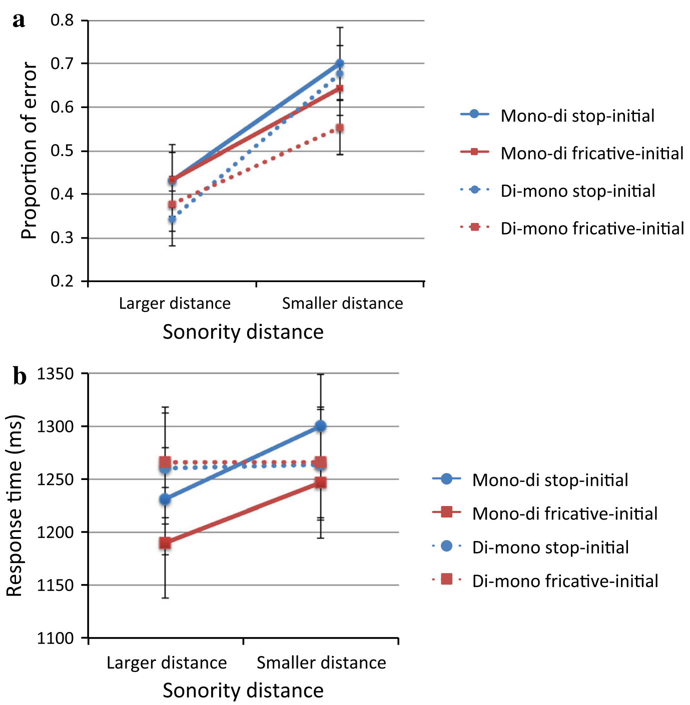


Fig. 2 Mean proportion of error (**a**) and mean response time (ms) (**b**) of the stop-initial and fricative-initial non-identical trials in Experiment 2 as a function of sonority distance. *Error bars* reflect the confidence intervals constructed for the difference among the means. ("Mono" = monosyllabic items, "di" = disyllabic items.)

an alternative explanation, misidentification stems from a failure to extract the phonetic form of the input from the acoustic input (Wright 2004). To adjudicate between these explanations, Experiment 3 investigates the identification of printed materials.

Past research shows that skilled readers assemble phonological representations during silent reading (e.g., Berent and Perfetti 1995; van Orden et al. 1990). Moreover, the phonological representation of printed words is shaped by phonological restrictions, including the grammatical restrictions on sonority (e.g., Berent and Lennertz 2010; Berent et al. 2009; Lennertz and Berent 2013). Accordingly, the grammatical repair hypothesis predicts that the difficulty in processing ill-formed onsets should persist even when presented in print.

Experiment 3 thus repeated the AX discrimination experiment using printed materials. As in Experiment 2, participants were asked to determine if the items that appear on the screen in succession are identical (*bzam-BZAM*) or not (*bzam-BEZAM*). To encourage phonological encoding, the two items were presented in different cases (e.g., *bdam-BEDAM*), and the SOA was increased from 1,500 to 2,500 ms. Because the printed modality inherently controls for the phonetic properties of stops and fricatives, we were now able to directly compare the best-formed sonority rise (e.g., *bzam*), sonority plateau (e.g., *bdam*) and sonority fall (e.g., *zbam*) in a three-way contrast (see Table 5). If small sonority distances are subject to grammatical repair, then as sonority distance decreases, participants should experience greater difficulty in discriminating monosyllables from their disyllabic counterparts. The replication of this finding with printed materials would rule out acoustic explanations for the results.

Table 5 The design of Experiment 3 (Printed stimuli)

Obstruent place	Sonority distance		
	Rise	Plateau	Fall
Labial-initial	*bzam*	*bdam*	*zbam*
Coronal-initial	*dvam*	*vzam*	*vdam*

Method

Participants

Thirty native English speakers, undergraduate students at Northeastern University, participated in this experiment in partial fulfillment of a course requirement. None of the students participated in Experiments 1 or 2. One of the subjects was excluded because his accuracy for both the monosyllabic identical (e.g., *bdam-bdam*) and non-identical (e.g., *bdam-bedam*) trials fell more than 2 standard deviations below the group mean.

Materials

The stimulus materials consisted of 72 printed monosyllables and 72 printed disyllables (see Appendix 2). Monosyllables were arranged in matched triplets, manifesting a large sonority rise (e.g., *bzam*), a sonority plateau (e.g., *bdam*) or a sonority fall (e.g., *zbam*).

Sonority rises were stop–fricative combinations; half were labial-initial (*bzam*); the other half were coronal-initial (*vdam*). Sonority falls were generated by reversing the order of consonants in their matched rises (e.g., *zbam, vdam*), whereas plateaus were invariably labial-initial. For sake of brevity, we refer to the triplets with labial-initial rises as "labial-initial" whereas those with coronal-initial rises are called "coronal-initial". Each labial-initial triplet was matched to a coronal-initial triplet for the rhyme (*bzam, bdam, zbam, dvam, vzam, vdam*).

The disyllabic items were created by inserting the letter *e* (or *E*) (*bzam* → *b<u>e</u>zam*). These items were arranged in two lists, balanced for identity, sonority distance, obstruent place and presentation order. Each list included 144 trials. Except for modality, the practice material was identical to that of Experiment 2.

Procedure

After initiating a trial, participants were presented with the first member of a stimulus pair in lower-case letters (e.g., *bdam*). The item remained on screen for 500 ms, and it was then replaced by a masking stimulus (XXXXXXX), displayed for 2,500 ms, followed by the second item (presented for 500 ms in upper-case letters (e.g., *BEDAM*). Participants were asked to judge whether the two items were identical (by pressing "1") or not (by pressing "2"). Participants were also given feedback on response time.

Prior to the experimental session, participants practiced the task using existing English words. During the practice, participants received feedback on both speed and accuracy. The order of the trials was randomized and the whole procedure took about 25 min.

Results

As in Experiment 2, identical (e.g., *bzam-BZAM*) and non-identical trials (e.g., *bzam-BEZAM*) were analyzed separately. Outliers amounted to 2 % of the data set.

Table 6 Mean response time in Experiment 3

Mean response time	Rise	Plateau	Fall
Identical trials			
Monosyllabic–disyllabic trials	635 (100)	638 (96)	628 (111)
Disyllabic–monosyllabic trials	656 (109)	666 (134)	663 (106)
Non-identical trials			
Monosyllabic–disyllabic trials	696 (122)	687 (120)	691 (125)
Disyllabic–monosyllabic trials	674 (108)	690 (116)	692 (104)

Standard deviations are indicated in parentheses

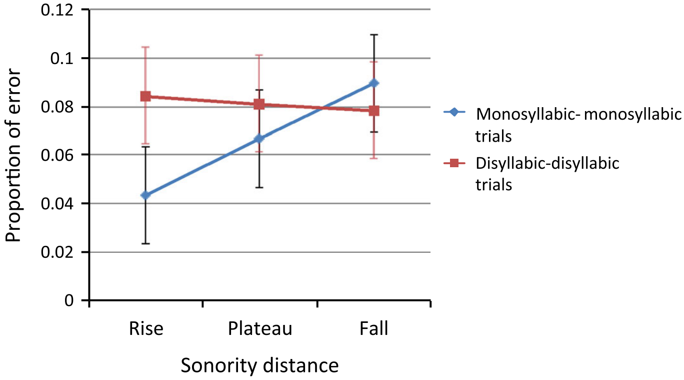


Fig. 3 Mean proportion of error of the stop-initial and fricative-initial identical trials in Experiment 3 as a function of sonority distance. *Error bars* reflect the confidence intervals constructed for the difference among the means

Identical Trials

We submitted the responses to identical trials (i.e., *bzam-BZAM, bezam-BEZAM*) to 2 SYLLABLE (i.e., monosyllabic–monosyllabic or disyllabic–disyllabic) × 2 OBSTRUENT PLACE (i.e., labial-initial vs. coronal-initial) × 3 SONORITY DISTANCE (i.e., rise vs. plateau vs. fall) ANOVAs. The interaction between syllable and sonority distance was marginally significant in the analyses of response accuracy ($F1(2, 56) = 2.91$, MSE = .01, $p < .06$, $F2(2, 22) = 3.16$, MSE = .003, $p < .09$; In response time: $F1(2, 56) = 7.93$, MSE = 9,505, $p < .01$, $F2(2, 22) = 1.99$, MSE = 14,577, $p < .19$, for the means, see Table 6).[8]

An inspection of the means (see Fig. 3) suggests that worse-formed monosyllables produced more errors than better-formed ones. The 2 OBSTRUENT PLACE (i.e., labial-initial vs. coronal-initial) × 3 SONORITY DISTANCE (i.e., rise vs. plateau vs. fall) ANOVAs on response accuracy to monosyllables indeed yielded a marginally significant effect of sonority distance ($F1(2, 56) = 2.64$, MSE = .01, $p < .08$; $F2(2, 22) = 4.46$, MSE = .003, $p < .02$).[9] Planned

[8] All accuracy results were supported by mixed-effect logit models (besides the sum coding of two-way contrasts syllable and obstruent type, we used forward difference coding for the three-way contrast sonority distance. Subject and sextet were included as random effects). The 2 syllable × 2 obstruent type × 3 sonority distance model confirmed yielded a marginal interaction of syllable and sonority distance ($\beta = 0.21786$, SE = 0.13129, Z = 1.659, $p < .097$).

[9] The mixed-effects 2 obstruent type × 3 sonority distance model yielded a marginally significant effect in the identical trials ($\beta = 0.3957$, SE = 0.2081, Z = 1.902, $p < .0572$), thus confirming the effect we found in the corresponding ANOVA.

Table 7 Mean proportion error to non-identical trials in Experiment 3

Mean proportion of error	Rise	Plateau	Fall
Monosyllabic–disyllabic trials	.16 (.20)	.13 (.15)	.12 (.15)
Disyllabic–monosyllabic trials	.08 (.12)	.07 (.11)	.09 (.12)

Standard deviations are indicated in parentheses

comparisons demonstrated that the worst-formed sonority fall produced reliably more errors than sonority rises (t1(56) = 2.3, $p < .03$, t2(22) = 2.85, $p < .01$). Responses to sonority plateaus did not reliably differ from either rises or falls.

Non-Identical Trials

Responses to non-identical items were submitted to 2 PRESENTATION ORDER (i.e., monosyllabic–disyllabic or disyllabic–monosyllabic) × 2 OBSTRUENT PLACE (i.e., labial-initial vs. coronal-initial) × 3 SONORITY DISTANCE (i.e., rise vs. plateau vs. fall) ANOVAs. The ANOVA yielded no main effect of sonority distance (both $F \leq 1.11$) or an interaction (all $F \leq 2.21$, $p \geq .12$) (see Table 7). However, the effect of presentation order was reliable (F1(1, 28) = 7.33, MSE = .04, $p < .01$, F2(1, 11) = 11.13, MSE = .01, $p < .01$). Participants were less accurate when monosyllables were followed by disyllables (e.g., *bzam-BEZAM*) relative to the opposite order (e.g., *bezam-BZAM*).

Similar ANOVAs on response time only yielded a marginally reliable effect of obstruent place (F1(1, 28) = 3.01, MSE = 5,616, $p < .09$, F2(1, 11) = 3.95, MSE = 1,289, $p < .07$), as labial-initial items produced faster responses than coronal-initial ones. No other effects were reliable ($F \leq 1.23$, $p \geq .29$) (see Table 6).

Discussion

The results of Experiment 3 suggest that English speakers remain sensitive to the minute sonority cline in obstruent–obstruent onsets even when they are presented in print. Identical items with sonority rise (e.g., *bzam-BZAM*) were identified more accurately than sonority falls (e.g., *zbam-ZBAM*). Although participants did not reliably differentiate the best- and worst-formed onsets from the intermediate sonority plateaus, responses to those items fell in between those two endpoints.

Note that, unlike auditory items, the effect of sonority with printed items obtained for the identity (as opposed to the nonidentity) trials. This difference might be due to the increase in the processing demands of identical printed items. Unlike the spoken items, printed identity trials consisted of two distinct tokens presented in different cases (e.g., *bdam-BDAM*), and the SOA was further increased in order to encourage the encoding of the first items. The elevated processing demands could have increased the reliance on phonological working memory, and consequently, monosyllables were now more vulnerable to repair (i.e., *bdam* → *bedam*). This explanation is indeed consistent with the observed order effect, whereby trials that required the maintenance of monosyllables in working memory produced more errors.

Crucially, the processing demands of monosyllables were modulated by their sonority distance. Items with small sonority distances (e.g., the sonority fall *zbam*) were more likely to undergo repair than those with large sonority distances (e.g., the sonority rise *bzam*). The replication of these results in the absence of any acoustic processing suggests that this effect might be due to the phonological structure of these items.

General Discussion

The present research investigated speakers' preferences concerning the sonority profile of stop–fricative onsets. Across languages, fricatives are more sonorous than stops. English, however, does not systematically enforce this distinction, as both types of obstruents are allowed in onset clusters (e.g., *pluck, flock*). Our research examined whether English speakers are nonetheless sensitive to the difference between the sonority levels of stops and fricatives.

If English speakers treat fricatives as more sonorous than stops, then they should be able to detect the slight sonority cline in onsets containing stops and fricatives. Stop–fricative onsets (e.g., *bzam, dvam*), should exhibit a small rise in sonority, so they should be preferred to the sonority plateaus in stop–stop and fricative–fricative onsets (e.g., *bdam* and *vzam*, respectively), which, in turn, should be favored to the sonority falls in fricative–stop onsets (e.g., *zbam, vdam*). And since the worse-formed small sonority distances are prone to grammatical repair (Berent et al. 2007, 2008, 2009, 2010, 2011a,b,c; Berent and Lennertz 2010; Lennertz and Berent 2013; Zhao and Berent 2013), we expect that small sonority clines should be misidentified less accurately than larger clines.

The results from Experiments 1–3 are consistent with this prediction. Experiment 1 showed that monosyllables with smaller sonority distances were prone to misidentification as disyllables (e.g., *bdam* was misidentified as *bedam* more often than *bzam* as *bezam*). Conversely, the disyllabic counterparts of worse-formed monosyllables (i.e., *bedam*, counterpart of *bdam*) were more readily identified, and this effect obtained even after controlling for duration of the intermediate vowel (i.e., the schwa, as in *bə̲dam*). The structure of stop–fricative onsets further modulated the discrimination of monosyllables from their disyllabic counterparts in the AX task (in Experiment 2). Here, we found that participants had greater difficulty with the trials worse-formed monosyllables (i.e., *bdam-bedam)* relative to better formed ones (e.g., *bzam-bezam*). Experiment 3 replicated the sensitivity to onset structure with printed materials. In particular, identical trials containing worse-formed sonority falls (i.e., *zbam-ZBAM*) were identified less accurately than better-formed sonority rises (i.e., *bzam-BZAM*).

The consistent difficulty in processing the minute sonority clines in stop–fricative combinations irrespective of input modality—for either spoken or printed materials—suggests some abstract knowledge that renders such onsets dispreferred. These results, however, do not speak to the origin of this knowledge. And indeed, the consistent preference for stop–fricative onsets might reflect not universal grammatical constraints, but rather the statistical similarity of these items to the English language. According to this statistical account, better-formed onsets are relatively immune to misidentification because they resemble existing English onsets more than worse-formed onsets do. For example, *bzam* might resemble existing English words more than *bdam* does.

To evaluate this possibility, we estimated the statistical similarity of our auditory and printed materials to English words. For auditory words, we calculated the number of phonological neighbors (i.e., existing words that were created by substituting a single phoneme), the neighbors' phonological frequency (i.e., summed frequency of neighbors), position-specific phoneme probability (i.e., the probability of a phoneme occurring in a given position, averaged across the four positions) and bi-phone probability of our auditory stimuli (i.e., the probability of two adjacent phonemes co-occurring in a given position, averaged across the biphones). For our printed materials (in Experiment 3), we likewise computed the number of orthographic neighbors (i.e., existing words that were created by substituting a single letter), the neighbors' frequency (i.e., summed frequency of neighbors), position-specific bigram count (i.e., the number of words sharing two adjacent letters in a specific position) and bigram frequency (i.e., the averaged frequency per million of the position-specific bigram

Table 8 Statistical properties of auditory materials in Experiments 1–2

	Auditory stimuli			
	Stop-initial items		Fricative-initial items	
	Better-formed (bz)	Worse-formed (bd)	Better-formed (vz)	Worse-formed (vd)
a. Number of neighbors	0.25	0.25	0	0
b. Neighbors' frequency (summed)	5	5	0	0
c. Position-specific phoneme probability	0.0269	0.0251	0.0197	0.0179
d. Bi-phone probability (summed)	0.0005	0.0004	0.0005	0.0004

Table 9 Statistical properties of the visual stimuli in Experiment 3

	Visual stimuli					
	Labial-initial items			Coronal-initial items		
	Rise (bz)	Plateau (bd)	Fall (zb)	Rise (dv)	Plateau (vz)	Fall (vd)
a. Number of neighbors	0.67	0.83	0.17	0.58	0.17	0.33
b. Neighbors' frequency (summed)	73.89	81.76	5.06	61.78	9.49	30.46
c. Bigram count	4.28	4.53	4.61	4.61	4.61	4.53
d. Bigram frequency (summed)	3,020.24	3,067.12	3,020.91	3,083.19	3,020.24	3,067.12

counts). We obtained the phoneme and bi-phone probabilities from the Phonotactic Probability Calculator (Vitevitch and Luce 2004), the phonological neighborhood measures were based on the Speech and Hearing Lab Neighborhood Database (Nusbaum et al. 1984) and the orthographic measures are based on the Orthographic Wordform Database (a CELEX-based corpus, Medler and Binder 2005).The summary statistics for auditory and printed materials are provided in Tables 8, 9, respectively.

An inspection of the means reveals no systematic correspondence between the statistical properties of the materials and their structure. For example, in the case of printed materials, better-formed labial onsets with rising sonority exhibited *lower* bigram frequency than worse formed onsets of level sonority.

To further assess whether statistical factors can explain our findings, we submitted the data into step-wise linear regression analyses. To determine the unique contribution of sonority distance, we first forced into the model the combined statistical factors (in step 1), followed by obstruent place (in step 2); sonority distance was entered last (in step 3). We also ran the reverse analyses to examine the unique effect of statistical factors (entered last). Since no effects of the response time models reached significance (all R^2 change $\leq .093$, $p \geq .152$), we only report the models using response accuracy. Results (see Table 10) showed that the statistical properties of the materials did not uniquely capture behavior in none of the three

Table 10 Step-wise linear regression analyses examining the contribution of statistical properties and sonority distance in Experiments 1–3

Experiment	Condition	Last predictor	Predictors forced in previous steps	R2 change	F change	*df*	*p* <
1	Non-identical trials	1. Sonority distance	Statistical properties, obstruent type	0.544	105.391	1, 41	0.001
		2. Statistical properties	Sonority distance, Obstruent type	0.06	2.902	4, 41	NS
2	Identical trials	1. Sonority distance	Statistical properties, obstruent type	0.043	2.36	1, 41	NS
		2. Statistical properties	Sonority distance, obstruent type	0.119	1.616	4, 41	NS
	Non-identical trials	3. Sonority distance	Statistical properties, obstruent type	0.242	13.839	1, 41	0.001
		4. Statistical properties	Sonority distance, obstruent type	0.04	0.567	4, 41	NS
3	Identical trials	1. Sonority distance	Statistical properties, obstruent place	0.063	5.047	1, 65	0.028
		2. Statistical properties	Sonority distance, obstruent place	0.043	0.855	4, 65	NS
	Non-identical trials	3. Sonority distance	Statistical properties, obstruent place	0.044	3.525	1, 65	NS
		4. Statistical properties	Sonority distance, obstruent place	0.099	1.979	4, 65	NS

experiments. In contrast, the unique contribution of sonority distance was found significant in all experiments, even after controlling for statistical similarity.

These analyses do not rule out all statistical explanation for the results. In particular, it is conceivable that the sensitivity to the sonority levels of stops and fricatives is learnable by more sophisticated grammatical models (e.g., the Maximum entropy model, Hayes and Wilson 2008). Nonetheless, our findings demonstrate a striking convergence between cross-linguistic preferences and the linguistic behavior of individual speakers. These results are consistent with the possibility of universal grammatical restrictions on the phonological system (e.g., Prince and Smolensky 1993/2004). The precise source for this convergence awaits further research.

Appendix 1

See Table 11.

Table 11 Monosyllabic items used in the auditory experiments (Experiments 1 and 2)

	Better-formed	Worse-formed
Stop-initial	bzim	bdim
	bzɑm	bdɑm
	bzɔm	bdɔm
	bzʊm	bdʊm
	bzin	bdin
	bzɑn	bdɑn
	bzɔn	bdɔn
	bzʊn	bdʊn
	bzig	bdig
	bzɑg	bdɑg
	bzɔg	bdɔg
	bzʊg	bdʊg
Fricative-initial	vzim	vdim
	vzɑm	vdɑm
	vzɔm	vdɔm
	vzʊm	vdʊm
	vzin	vdin
	vzɑn	vdɑn
	vzɔn	vdɔn
	vzun	vdun
	vzig	vdig
	vzɑg	vdɑg
	vzɔg	vdɔg
	vzʊg	vdʊg

Appendix 2

See Table 12.

Table 12 Monosyllabic items used in the printed experiment (Experiment 3)

	Sonority rise	Sonority plateau	Sonority fall
Labial-initial	bzim	bdim	zbim
	bzam	bdam	zbam
	bzom	bdom	zbom
	bzum	bdum	zbum
	bzin	bdin	zbin
	bzan	bdan	zban
	bzon	bdon	zbon
	bzun	bdun	zbun
	bzig	bdig	zbig
	bzag	bdag	zbag
	bzog	bdog	zbog
	bzug	bdug	zbug
Coronal-initial	dvim	vzim	vdim
	dvam	vzam	vdam
	dvom	vzom	vdom
	dvum	vzum	vdum
	dvin	vzin	vdin
	dvan	vzan	vdan
	dvon	vzon	vdon
	dvun	vzun	vdun
	dvig	vzig	vdig
	dvag	vzag	vdag
	dvog	vzog	vdog
	dvug	vzug	vdug

References

Barlow, J. (2005). Sonority effects in the production of consonant clusters by Spanish-speaking children. In *Paper presented at the selected proceedings of the 6th conference on the acquisition of Spanish and Portuguese as first and second languages*.

Barlow, J. (2001). A preliminary typology of initial clusters in acquisition. *Clinical Linguistics & Phonetics*, *15*(1/2), 9–13.

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, *102*(1), 146–184.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. [Research Support, N.I.H., Extramural]. *Cognition*, *104*(3), 591–630.

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*(14), 5321–5325.

Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, *26*(01), 75–108.

Berent, I., & Lennertz, T. (2010). Universal constraints on the sound structure of language: Phonological or acoustic? *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 212–223.

Berent, I., Balaban, E., Lennertz, T., & Vaknin-Nusbaum, V. (2010). Phonological universals constrain the processing of nonspeech stimuli. *Journal of Experimental Psychology: General*, *139*(3), 418–435. doi:10.1037/a0020094.

Berent, I., Balaban, E., & Vaknin-Nusbaum, V. (2011a). How linguistic chickens help spot spoken-eggs: Phonological constraints on speech identification. *Frontier in Language Sciences*, *2*, doi:10.3389/fpsyg.2011.00182.

Berent, I., Harder, K., & Lennertz, T. (2011b). Phonological universals in early childhood: Evidence from sonority restrictions. *Language Acquisition*, *18*(4), 281–293.

Berent, I., Lennertz, T., & Balaban, E. (2011c). Language universals and misidentification: A two way street. *Language and Speech*, *55*, 1–20.

Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics*, *32*(2), 117–166.

Boersma, P., Weenink, D. (2003). Praat: Doing phonetics by computer (version 5.2.32). Retrieved from http://www.praat.org/

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, *7*, 35–59.

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*, 381–410.

Byrd, D. (1992). Perception of assimilation in consonant clusters: A gestural model. *Phonetica*, *49*, 1–24.

Christman, S. H. (1992). Uncovering phonological regularity in neologisms: Contributions of sonority theory. *Clinical Linguistics & Phonetics*, *6*, 219–247.

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 physics of speech* (pp. 282–333). Cambridge: Cambridge University Press.

Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, *38*(2), 272–288. doi:10.1016/j.wocn.2010.01.001.

Davidson, L. (2011a). Phonetic and phonological factors in the second language production of phonemes and phonotactics. *Language and Linguistics Compass*, *5*(3), 126–139.

Davidson, L. (2011b). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 270–282.

Davidson, L. (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics*, *40*(2), 234–248.

Dell, F., & Elmedlaoui, M. (1985). Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics*, *7*, 105–130.

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*(6), 1568–1578.

Dupoux, E., Parlato, E., Frota, S., Hirose, Y., & Peperkamp, S. (2011). Where do illusory vowels come from? *Journal of Memory and Language*, *63*(3), 199–210.

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.

Fleischhacker, H. A. (2005). *Similarity in phonology: Evidence from reduplication and loan adaptation. PhD*. Los Angeles: University of California.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language & Hearing Research*, *42*(3), 708–726.

Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition*. Cambridge: Cambridge University Press.

Gouskova, M. (2001). Falling sonority onsets, loanwords, and syllable contact. CLS 37: The main session. In *Papers from the 37th meeting of the Chicago linguistic society* (pp. 175–185).

Greenberg, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word, 20*, 157–177.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. *Universals in Human Language*, *2*, 243–279.

Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, *39*, 379–440.

Klippenstein, R. (2008). Word-initial consonant clusters in Albanian. *Ohio State University Working Papers in Linguistics, 59*, 10–32.

Ladefoged, P. (2001). *A course in phonetics*. Boston: Heinle & Heinle.

Lennertz, T., & Berent, I. (2013). *People's knowledge of phonological universals: Evidence from fricatives and stops*. manuscript submitted for publication.

Massaro, D. W., & Cohen, M. M. (1983). Phonological constraints in speech perception. *Perception & Psychophysics*, *34*, 338–348.

McCarthy, J. J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, *17*, 207–263.

Medler, D. A., & Binder, J. R. (2005). *MCWord: An on-line orthographic database of the English language*.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*(1), 55–71.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report*, *10*, 357–376.

Ohala, D. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders*, *32*, 397–422.

Parker, S. (2002). *Quantifying the sonority hierarchy*. PhD, University of Massachusetts.

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, *36*, 55–90.

Pertz, D. L., & Bever, T. G. (1975). Sensitivity to phonological universals in children and adolescents. *Language*, *51*, 149–162.

Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics*, *60*, 941–951.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. (Version 2.10.1). Vienna, Austria. Retrieved from http://www.R-project.org

Redford, M. A. (2008). Production constraints on learning novel onset phonotactics. *Cognition*, *107*, 785–816.

Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language*, *64*, 83–121.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh PA: Psychology Software Tools.

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.

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 (Linguistic and philosophical implications)* (Vol. 2, pp. 27–160). Cambridge, MA: 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.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. PhD, MIT, Cambridge, MA.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–46.

van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488–522.

Vitevich, M., & Luce, P. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *40*, 374–408.

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.

Wright, R. (2004). A review of perceptual cues and cue robustness. In D. Steriade, R. M. Kirchner, & B. Hayes (Eds.), *Phonetically based phonology*. Cambridge: Cambridge Univ Press.

Yavas, M., & Gogate, L. (1999). Phoneme awareness in children: A function of sonority. *Journal of Psycholinguistic Research*, *28*, 245–259.

Zhao, X., & Berent, I. (2013). *Are markedness constraints universal? Evidence from Mandarin Chinese speakers*. Manuscript submitted for publication.