J Psycholinguist Res
DOI 10.1007/s10936-017-9510-2

The Basis of the Syllable Hierarchy: Articulatory Pressures or Universal Phonological Constraints?

Xu Zhao[1] · **Iris Berent**[1]

Abstract Across languages, certain syllable types are systematically preferred to others (e.g., *blif* $\succ$ *bnif* $\succ$ *bdif* $\succ$ *lbif*, where $\succ$ indicates a preference). Previous research has shown that these preferences are active in the brains of individual speakers, they are evident even when none of these syllable types exists in participants' language, and even when the stimuli are presented in print. These results suggest that the syllable hierarchy cannot be reduced to either lexical or auditory/phonetic pressures. Here, we examine whether the syllable hierarchy is due to articulatory pressures. According to the *motor embodiment* view, the perception of a linguistic stimulus requires simulating its production; dispreferred syllables (e.g., *lbif*) are universally disliked because their production is harder to simulate. To address this possibility, we assessed syllable preferences while articulation was mechanically suppressed. Our four experiments each found significant effects of suppression. Remarkably, people remained sensitive to the syllable hierarchy regardless of suppression. Specifically, results with auditory materials (Experiments 1–2) showed strong effects of syllable structure irrespective of suppression. Moreover, syllable structure uniquely accounted for listeners' behavior even when controlling for several phonetic characteristics of our auditory materials. Results with printed stimuli (Experiments 3–4) were more complex, as participants in these experiments relied on both phonological and graphemic information. Nonetheless, readers were sensitive to most of the syllable hierarchy (e.g., *blif* $\succ$ *bnif* $\succ$ *bdif*), and these preferences emerged when articulation was suppressed, and even when the statistical properties of our materials were controlled via a regression analysis. Together, these findings indicate that speakers possess broad grammatical preferences that are irreducible to either sensory or motor factors.

Keywords Phonology · Universal grammar · Embodiment · Articulation

✉ Iris Berent
i.berent@neu.edu

1 Department of Psychology, Northeastern University, 125 Nightingale, 360 Huntington Ave, Boston, MA 02115, USA

Published online: 14 July 2017

Introduction

Although natural languages vary in many ways, they nonetheless converge on certain aspects of their design. Consider, for example, the structure of onset clusters (i.e., the consonant sequence at the beginning of a syllable, e.g., ***bl**ack*). Across languages, onsets like *bl* are preferred to *lb* (i.e., *blif* $\succ$ *lbif*, "$\succ$" indicates preference; Greenberg 1978). Moreover, languages that tolerate onsets such as *lb* tend to allow onsets like *bl*. Such observations suggest that speakers of different languages share common restrictions on language structure.

The nature of these constraints, however, remains unclear. One explanation attributes phonological regularities to the *grammar*—a set of abstract phonological principles. For example, Optimality Theory (e.g., Prince and Smolensky 1993/2004) asserts that all grammars share universal constraints on syllable structure. Furthermore, these constraints are active in the brains of all speakers, irrespective of whether these syllables are present or absent in their language. Structures that abide by these constraints (e.g., *blif*) are relatively well-formed, hence, they are preferred to those that violate them (e.g., *lbif*).

But on alternative *non-grammatical* explanations, the restrictions on syllable structure emanate only from nonlinguistic sources. A lexical account asserts that syllables like *blif* are preferred because they are similar to familiar words (e.g., *blif* is similar to *block* and *sniff*). On a phonetic explanation, *lbif* is only disliked because people fail to encode its phonetic form from the speech input (Blevins 2004; Bybee 2008; Evans and Levinson 2009).

Past research has amassed multiple challenges against each of these two explanations (evidence we review below). The existing literature, however, leaves open the possibility that the dislike of ill-formed syllables is due to motor pressures. According to the motor account, the perception of a word requires covert simulation of its articulatory production. For example, upon hearing *lbif*, people attempt to covertly produce it. The dislike of *lbif*, then, could be due to the difficulty in its covert simulation.

Our present research examines this possibility. We begin by outlining the grammatical explanation for syllable structure, and reviewing the relevant experimental evidence. A description of our present manipulations and results follows.

The Syllable Hierarchy and Its Sources: Grammatical Versus Non-grammatical Explanations

A prominent linguistic account attributes the restrictions on syllable structure to sonority. Sonority (s) is an abstract phonological property of segments that correlates with acoustic intensity (Clements 2005; Parker 2008). Least sonorous are stops (e.g., b,p, $s = 1$), followed by fricatives (e.g., f,v, $s = 2$), nasals (e.g., m,n, $s = 3$), liquids (e.g., l,r, $s = 4$), and glides (e.g., w,y, $s = 5$). Accordingly, onsets such as *bl* exhibit a large rise in sonority ($\Delta s = s(l) - s(b) = 3$), *bn* manifests a small rise ($\Delta s = 2$), *bd* levels in sonority ($\Delta s = 0$) and *lb* falls in sonority ($\Delta s = -3$). Since the grammar bans small sonority distances, syllables with small sonority distances are avoided—the smaller the distance, the stronger its dislike. Together, the restriction on sonority distance gives rise to the following ***syllable hierarchy*** (e.g., *bl* $\succ$ *bn* $\succ$ *bd* $\succ$ *lb*, where $\succ$ indicates a preference).

To test this *grammatical* explanation, our past research examined whether speakers are sensitive to this hierarchy (along with related generalizations). Sensitivity to syllable structure was inferred from its effect on perception. We reasoned that, if the grammar bans onsets with small sonority distances, then ill-formed syllables will not be encoded faithfully by the grammar. Instead, ill-formed syllables (e.g., *lbif*) will be repaired by the grammar (e.g., by inserting a schwa in between the consonants of the onset cluster, *lbif* → *lebif*, etc.). As a result, speakers will be more likely to misidentify ill-formed syllables (e.g., *lbif*) compared to well-formed ones (e.g., *blif*)—the worse formed a syllable, the more likely its misidentification.

This prediction was borne out by the results of multiple experiments: as the syllable becomes worse formed on the hierarchy, misidentification increases monotonically. Critically, speakers are demonstrably sensitive to the syllable hierarchy even when most syllable types do not exist in their language (e.g., English: Berent et al. 2007; Spanish: Berent et al. 2012; French: Maïonchi-Pino et al. 2012; Hebrew: Berent et al. 2013), and even when their language does not allow any onset clusters at all (e.g., Korean: Berent et al. 2008; Mandarin Chinese: Zhao and Berent 2015). In fact, recent findings have documented sensitivity to the hierarchy even in neonates (Gómez et al. 2014).

Additional results suggest that onset preference is not solely due to auditory/phonetic failures to encode the acoustic input (Berent et al. 2007, 2012). Indeed, the misidentification of ill-formed structures has been found even when the materials were presented visually, as printed words (Berent 2008; Berent et al. 2009; Berent and Lennertz 2010; Tamási and Berent 2014). Together, these results appear to demonstrate that people converge on shared phonological preferences despite minimal linguistic experience with the relevant syllables, and irrespective of the input modality of these stimuli (i.e., auditory or visual). These findings are in line with the possibility that people share universal grammatical constraints on language structure.

These results, however, leave open the possibility that the syllable hierarchy results from articulatory demands. According to the *motor embodiment* account,[1] the perception of spoken words requires the hearer to simulate their production by means of sub-vocal articulation (e.g., Lakoff and Johnson 1999; MacNeilage 2008). Extending this explanation to the syllable hierarchy, this account would assert that the ban on ill-formed syllables (e.g., *lbif*) reflects not universal linguistic bans, but rather the difficulties in their covert sub-vocal simulation. In what follows, we review some of the evidence in support of this theory.

The Role of the Articulatory Motor System in Speech Perception

A large literature suggests that articulatory action constrains speech perception. One source of evidence in support for the *embodiment* account is presented by neurophysiological studies. Using fMRI, Pulvermüller et al. (2006) found that listening to speech sounds activates the motor representations of their articulators: labial sounds (e.g., /*p*/) activate lip motor sites relative to the motor representation of the tongue, whereas sounds produced by the tongue (e.g., the coronal /*t*/) trigger the opposite pattern. Similarly, Fadiga et al. (2002) showed that passive listening to coronal sounds (e.g., /*t*/) engaged the tongue-related motor centers and induced motor-evoked potentials in the tongue muscle.

Other neurophysiological evidence suggests that the activation of these motor sites plays a causal role in speech perception. These findings were obtained using a transcranial magnetic stimulation (TMS) methodology. TMS is a noninvasive method that temporarily modulates activity in targeted brain regions by delivering electromagnetic pulses to the cortical surface

[1] Our present discussion focuses on a strong version of the embodied motor theory of speech perception. According to that view, speech perception requires actual articulatory action (e.g., Schwartz et al. 2002). Other motor theories [e.g., Motor Theory by Liberman and Mattingly (1985) and the Direct Realist Theory (Fowler 1986)] allow for at least some abstraction of motor gestures (for reviews, see Diehl et al. 2004; Galantucci et al. 2006; Samuel 2011; Schwartz et al. 2002).

(O'Shea and Walsh 2007). Results from the TMS literature show that the perception of speech sounds is selectively modulated by the stimulation of the relevant articulator. For example, stimulating the tongue motor region facilitates the perception of coronal phonemes (e.g., /*t*/ and /*d*/) whereas the stimulation of the lip motor area facilitates the perception of labial phonemes (e.g., /*p*/ and /*b*/) (D'Ausilio et al. 2009, 2012). Similarly, the stimulation of the lip motor area impaired the identification labial sounds like /*p*/ relative to coronal sounds like /*t*/ (Möttönen and Watkins 2009).

While these studies demonstrate that articulatory motor regions play a causal role in ***phonetic*** categorization (e.g., /*p*/ vs. /*t*/), it is unclear whether motor activation (and furthermore, motor simulation) is also necessary for the ***phonological*** patterning of these sounds into syllables (e.g., /*p*/+/*l*/+/*a*/ vs. /*l*/+/*p*/+/*a*/). A large literature suggests that phonetic and phonological representations are distinct, inasmuch as phonological representations are discrete and algebraic, whereas phonetic representations are analog (e.g., Abler 1989; Berent 2013; Chomsky and Halle 1968). If phonetic and phonological representations are distinct, then the effects of TMS on phonetic categorization may not necessarily apply to phonology.

To examine this possibility, a recent TMS study (Berent et al. 2015) asked whether sensitivity to the syllable hierarchy (i.e., $bl \succ bn \succ bd \succ lb$) requires motor simulation. Participants in this experiment performed a syllable count task (e.g., does *lbif* include one syllable or two?), while undergoing stimulation of the lip motor area by TMS. The effect of TMS was compared to a Sham condition. If sensitivity to the syllable hierarchy requires motor simulation, then the disruption of the motor system by TMS should attenuate sensitivity to the syllable hierarchy. Furthermore, if the misidentification of ill-formed syllables is due to difficulties in their simulation, then the worst-formed onsets like *lb* should impose the greatest articulatory demands on the motor system. Consequently, the disruption of motor regions by TMS should exert the strongest effect (i.e. the strongest attenuation in the discriminate one syllable from two) with the *lb*-type.

Results countered this prediction. While the disruption of the lip motor area did attenuate overall performance in this task, participants remained sensitive to the syllable hierarchy (i.e., $bl \succ bn \succ bd \succ lb$) even when the motor area was disrupted by TMS. Furthermore, the effect of TMS was found not with the worst formed-syllables like *lbif*, but rather with the better-formed syllables (e.g., *bl* and *bn*). A subsequent fMRI study (Experiment 3, Berent et al. 2015; see also Berent et al. 2014) further showed that the processing of the worst formed *lb*-type syllables disengaged (rather than engaged the sensorimotor lip regions, predicted by the *embodiment* hypothesis, Berent et al. 2014). Together, these findings challenge the causal role of articulatory system in the computations of sound structures.

These TMS results (Berent et al. 2015), however, nonetheless suffer from several limitations. First, since TMS pulses typically reach only surface cortical regions (Bolognini and Ro 2010) and they only target a single articulator, this manipulation is unlikely to block motor simulation entirely. Moreover, it is unclear whether the effects of TMS are due to the disruption of motor simulation, or to other non-motor (e.g., phonological) functions. The finding that motor areas contribute to articulatory production does not rule out the possibility that these areas can also engage in other (non-motoric) functions [e.g., elementary memory functions (Sanes and Donoghue 2000); application of previously learned logic rules (Acuna et al. 2002); cutaneous and proprioceptive responses (Hatsopoulos and Suminski 2011)].

To overcome these limitations, the current study seeks to complement the previous TMS findings by studies that utilize a mechanical suppression of the motor system. Participants in our experiments were instructed to accommodate two tongue depressors in their mouth (one above and another below the tongue) in order to suppress tongue movement. In addition, participants were asked to point the two tongue depressors at the same direction, an action

Fig. 1 Demonstration of the control (*left panel*) and the suppression (*right panel*) condition

that also engaged the lip (for an illustration, see right panel of Fig. 1). Accordingly, the suppression method reduced activity in both the lip and tongue.

The *motor embodiment* account predicts that articulatory mechanical suppression should modulate the processing of linguistic stimuli (e.g., Glenberg et al. 2005). A recent study provides specific evidence for the effect of static suppression on speech perception (Bruderer et al. 2015). Bruderer and colleagues investigated the capacity of 6-month olds to discriminate two non-native sounds that contrast on the placement of the tongue tip. Results showed that, when the relevant articulator (i.e., tongue) was selectively suppressed (by a teething toy that blocked the tongue's movement), infants were no longer able to discriminate these two sounds. Together, these results demonstrate that mechanical suppression of motor articulators modulates speech perception. Building on this research, our present research uses mechanical suppression in order to investigate the nature of phonological restrictions.

The Present Research

This research examines the causal role of motor simulation in the computation of phonological structure. Specifically, we ask whether sensitivity to the syllable hierarchy (i.e., $bl \succ bn \succ bd \succ lb$) is due to grammatical principles or the demands of articulatory simulation. To address this question, we examine whether the effect of syllable structure is maintained when articulatory action is suppressed mechanically.

We approach this problem in two steps. First, we ask whether the suppression manipulation is effective, that is, whether suppression diminishes participant's overall ability to differentiate monosyllables from their disyllabic counterparts. Insofar as suppression effects are found, we next ask whether suppression further attenuates people's sensitivity to the syllable hierarchy ($blif \succ bnif \succ bdif \succ lbif$). The *grammatical* and *motor embodiment* account both predict that, as the syllable becomes worse formed, misidentification in the control condition should increase. The two accounts differ on their predictions regarding the effect of suppression.

If the syllable hierarchy is due to grammatical constraints (as predicted by the *grammatical* account), then sensitivity to the syllable hierarchy should obtain regardless of whether participants are free to engage in articulation. In contrast, if the misidentification of ill-formed syllables results from difficulties in their subvocal articulation, then articulatory suppression should attenuate the overall sensitivity to syllable structure (e.g., decrease the disadvantage of *lbif* relative to *blif*). And since worse formed syllables presumably impose greater articula-

Table 1 General procedure of suppression manipulation in Experiments 1–4

	List 1	List 2
Experimental block		
First block	Control (no suppression)	Suppression
Second block	Suppression	Control (no suppression)

All participants completed two experimental blocks; half followed the order specified in list 1 and the other half followed list 2. (see "Appendix A")

tory demands, they should be more susceptible to suppression than better-formed syllables. In fact, suppression should alleviate these articulatory difficulties, so it should potentially ***improve*** the identification of ill-formed onsets (i.e., more accurate performance for *lbif* in the suppression relative to the control condition).

To adjudicate between these possibilities, four experiments were conducted: Experiments 1–2 used auditory stimuli; in Experiment 1 people performed a syllable count task (e.g., does *lbif* has one syllable or two?) whereas Experiment 2 used an identity discrimination task (e.g., is *blif* identical to *belif*?). To address the possibility that the findings with auditory materials only reflect auditory/phonetic factors, Experiments 3–4 extended our investigation to printed materials, either with the absence or presence of background visual noise, respectively.

Experiment 1: Syllable Count

Experiment 1 examined the effect of articulatory suppression using a syllable count task. On each trial, participants heard a single nonword stimulus (either monosyllabic or disyllabic) and they were instructed to indicate whether this item had one syllable or two. Since the syllable task elicits a forced choice discrimination, we gauged performance using the sensitivity measure, d' (correct monosyllabic responses are considered 'hits').

To examine the possibility of carryover effects across the suppression conditions, in this and all subsequent experiments, each participant completed two experimental blocks. Half of the participants received the control condition (no articulatory suppression) in the first block, followed by articulatory suppression in the second block (they completed Experiment List 1); the other half were assigned to the reversed order of conditions (i.e., List 2; see Table 1). Accordingly, the list factor forms part of our analyses.

Method

Participants

Forty native English speakers, students of Northeastern University, took part in this experiment. In this and all subsequent experiments, participants received course credit for their participation. The procedures in this and all subsequent experiments were approved by the Institutional Review Board at Northeastern University. All participants signed an informed consent.

Materials

The materials consisted of pairs of monosyllabic nonwords and their matched disyllabic counterparts, described in Berent et al. (2007). Briefly, monosyllables were arranged in quartets whose onsets exhibited either large sonority rises, small rises, plateaus or falls in sonority (e.g., /trap/, /tmap/, /tpap/, /rtap/, respectively, see "Appendix B"). Disyllables differed from monosyllables by a schwa (e.g., */t∂rap/*, */t∂map/, /t∂pap/, /r∂tap/*). The materials included a total of 240 items (2 syllable: monosyllable/disyllable x 4 onset type: large rise/small rise/plateau/fall × 30 quartets), divided into two halves, matched for the number of onset type × syllable combination. These two halves were treated as two experimental sublists, and each such list was presented in a separate experimental block (with order counter-balanced), such that each participant completed all 240 trials (with 120 trials per block, see "Appendix A"). All items were recorded by a native Russian speaker (Russian allows all these onset types, so these items can be all produced naturally by Russian speakers).

Procedure

In this and all subsequent experiments, participants were randomly assigned to one of two experimental lists that contrasted on the block order of the suppression condition (control-suppression versus suppression-control, for lists 1 and 2, respectively). In the suppression condition, participants were instructed to safely accommodate two tongue depressors in their mouth pointing at the same direction—one above and the other beneath their tongue. In the control condition, participants performed the task normally, without tongue depressors. Trial order was randomized.

In all four experiments, participants were first familiarized with the task in a practice session with existing English words (e.g., *sport, support*), which preceded the experimental session. Slow responses (response time over 2500 ms) triggered a computerized warning message ("Too Slow!").

On each trial, participants were presented with a single auditory nonword. They were asked to quickly indicate the number of syllables by pressing one of two keys (1 = one-syllable; 2 = two-syllable).

Results

The effect of suppression on sensitivity (d') was examined using 2 suppression (suppression/control) ×4 onset type (large rise/small rise/plateau/fall)×2 list (control−suppression/ suppression−control) ANOVAs, conducted using participants (F1) and items (F2) as random variables.[2] Given near-floor performance with the worst-formed syllables, we did not conduct an analysis of response time.

The 3-way suppression × onset type × list interaction was marginally significant ($F1(3, 114) = 2.35, p = 0.076$; $F2(3, 174) = 2.34, p = 0.075$). This interaction suggests that the effect of suppression varied depending on the syllable type and the block order of the suppression manipulation in the experiment (recall that the list factor is closely linked to block order, see Table 1 for details).

To further examine the effect of suppression, we next compared the suppression and control conditions when they are both administered in the first versus second blocks of trials, separately. Our analyses examined two questions: (a) Does suppression modulate performance? and (b) Does suppression attenuate sensitivity to the syllable hierarchy?

a. Does suppression modulate performance? Figure 2 plots sensitivity (d') in the first (Fig. 2a) and second block (Fig. 2b) of trials (the accuracy and response time means are pre-

[2] The ANOVA was conducted because of its resilience to moderate deviations from normality (Glass et al. 1972; Martin and Games 1977; Harwell et al. 1992; Lix et al. 1996).

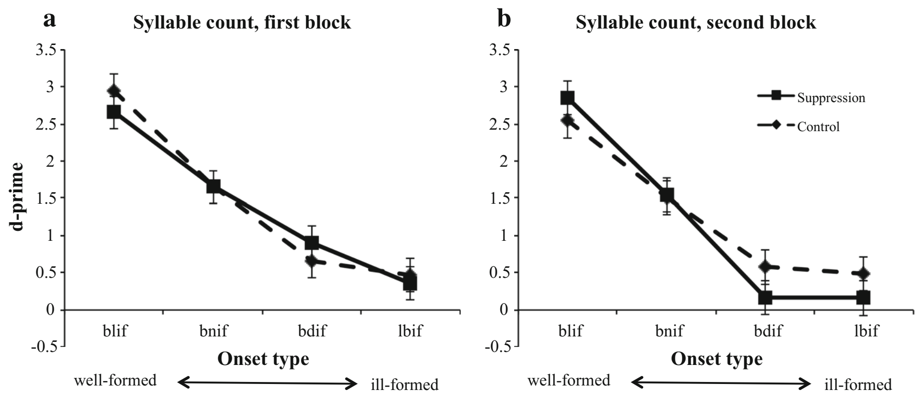


Fig. 2 The effect of articulatory suppression on sensitivity (d') to onset type in the first (**a**) and second block (**b**) of Experiment 1. *Note error bars* indicate 95% confidence intervals for the difference between the means

Table 2 Mean response accuracy (ACC, proportion correct) and correct response time (RT) in Experiment 1

	Monosyllabic items				Disyllabic items			
	Large rise	Small rise	Plateau	Fall	Large rise	Small rise	Plateau	Fall
ACC								
First block								
Suppression	0.93	0.56	0.32	0.17	0.83	0.89	0.92	0.92
Control	0.89	0.50	0.18	0.12	0.95	0.95	0.96	0.98
Second block								
Suppression	0.88	0.42	0.05	0.05	0.95	0.96	0.97	0.98
Control	0.90	0.52	0.26	0.18	0.84	0.89	0.89	0.92
RT (ms)								
First block								
Suppression	1038	1131	1215	1281	1116	1111	1121	1179
Control	960	1039	1153	1160	1038	1026	1035	1071
Second block								
Suppression	918	1033	1169	1180	928	933	921	961
Control	971	1071	1115	1064	1064	1063	1031	1060

sented in Table 2). An inspection of the means suggests that suppression affected sensitivity only in the second block.

The 2 suppression × 4 onset type ANOVAs in the first block did not yield a significant interaction (both $p > 0.13$) or a significant main effect of suppression (both $p > 0.14$). By contrast, in the second block, the suppression × onset type interaction was significant ($F1(3, 114) = 3.53$, $p < 0.018$; $F2(3174) = 4.23$, $p < 0.007$).

To interpret the interaction, we next compared responses to the suppression and control conditions for each of the four types of onsets presented in the second block. Planned comparisons (using participants, $t1$, and items, $t2$, as random variables) revealed that articulatory suppression *impaired* responses to onsets of level sonority (e.g., *bdif*; $t1(20) =$

1.77, $p = 0.079$; $t2(30) = 2.11$, $p < 0.036$). In contrast, for onsets with large rises (e.g., *blif*), suppression tended to *improve* performance, and this effect was marginally significant ($t2(30) = 2.12$, $p < 0.04$; albeit not by participant, $t1(20) = 1.31$, $p = 0.19$). For the remaining two types of onsets (e.g., *bnif, lbif*), the effect of suppression was not significant (all $p > 0.13$).

b. Does suppression attenuate sensitivity to the syllable hierarchy? Having shown that the suppression manipulation is effective, we next asked whether it attenuates sensitivity to syllable structure. An inspection of the means indicates that, as the onset became worse formed, sensitivity (d') decreased, and this trend emerged regardless of whether suppression was present or absent (Fig. 2).

The ANOVA (2 suppression $\times$ 4 onset type $\times$ 2 list) results were in line with this conclusion. In the control condition, the simple main effect of onset type was found significant in both the first ($F1(3, 57) = 112.02$, $p < 0.001$; $F2(3, 87) = 48.67$, $p < 0.001$) and second block ($F1(3, 57) = 45.89$, $p < 0.001$; $F2(3, 87) = 37.85$, $p < 0.001$). Planned comparisons revealed that onsets with large rises (e.g., *blif*) elicited better sensitivity (d') than small rises (e.g., *bnif*) in both blocks (first block: $t1(57) = 8.52$, $p < 0.001$; $t2(87) = 5.43$, $p < 0.001$; second block: $t1(57) = 5.14$, $p < 0.001$; $t2(87) = 5.11$, $p < 0.001$). Small rises, in turn, produced better sensitivity than plateaus (e.g., *bdif*; first block: $t1(57) = 6.57$, $p < 0.001$; $t2(87) = 4.61$, $p < 0.001$; second block: $t1(57) = 4.65$, $p < 0.001$; $t2(87) = 3.76$, $p < 0.001$). Responses to sonority plateaus and falls (e.g., *bdif* vs. *lbif*) did not differ significantly (first block: $t1(57) = 1.23$, $p = 0.22$; $t2(87) = 0.59$, $p = 0.55$; second block: $t1(57) = 0.45$, $p = 0.65$; $t2(87) = 0.57$, $p = 0.57$). Thus, as sonority distance decreased, participants in the control condition became less sensitive to onset structure (i.e., $blif \succ bnif \succ \{bdif, lbif\}$).

Crucially, the effect of onset type remained significant even when articulation was suppressed. Specifically, the simple main effect of onset type was significant in both the first ($F1(3, 57) = 39.19$, $p < 0.001$; $F2(3, 87) = 40.01$, $p < 0.001$) and second block ($F1(3, 57) = 146.87$, $p < 0.001$; $F2(3, 87) = 66.74$, $p < 0.001$). Planned comparisons showed that as sonority distance decreased, so did participants' sensitivity. Specifically, sensitivity to onsets with large rises (e.g., *blif*) was reliably better relative to small rises (e.g., *bnif*; first block: $t1(57) = 4.47$, $p < 0.001$; $t2(87) = 4.85$, $p < 0.001$; second block: $t1(57) = 8.66$, $p < 0.001$; $t2(87) = 5.93$, $p < 0.001$). Small rises, in turn, elicited better sensitivity than plateaus (e.g., *bdif*; first block: $t1(57) = 3.33$, $p < 0.002$; $t2(87) = 2.70$, $p < 0.009$; second block: $t1(57) = 9.16$, $p < 0.001$; $t2(87) = 6.17$, $p < 0.001$). Finally, the worst-formed onsets of falling sonority (e.g., *lbif*) produced even worse sensitivity compared to onsets with level sonority (e.g., *bdif*) in the first block of trials ($t1(57) = 2.43$, $p < 0.02$; $t2(87) = 2.98$, $p < 0.005$), albeit not in the second block ($t1(57) = 0.04$, $p = 0.96$; $t2(87) = 0.09$, $p = 0.93$). Thus, regardless of suppression, the effect of sonority distance emerged—as the onset became worse-formed, sensitivity decreased (i.e., $blif \succ bnif \succ bdif \succ lbif$).

Discussion

Experiment 1 demonstrates that articulatory suppression modulates sensitivity to syllable structure. The direction of this effect, however, stands in stark contrast to the *motor embodiment* hypothesis. While the *motor embodiment* hypothesis predicts that suppression should improve the identification of ill-formed syllables (as it releases these structures from the cost of erroneous simulation), we observed *impairment*, and this effect emerged only in the second block of trials. Moreover, suppression did not systematically attenuate speakers' sensitivity to the syllable hierarchy. Rather, when suppression was administered, the worst formed *lbif*-type syllables produced worse performed than the *blif* and *bnif*-type items. Although the contrast relative to *bdif* was not significant in the second block of trials, this null results is most likely due to a floor effect. Indeed, the overall effect of well-formedness, as measured by the differential responses to the best- compared to the worst-formed onsets (i.e., *blif–lbif*), was numerically *larger* in the suppression condition (Δd'=d'(*blif*)-d'(*lbif*) = 2.85–0.16 = 2.69) compared to the control condition (Δd' = 2.54–0.49 = 2.05). These findings counter the *motor embodiment* account.

Our results cannot clearly establish why the effect of suppression was selective to the second block of trials. One possibility is that this outcome is due to the attention demands of the suppression task, combined with fatigue, which presumably exacerbated those attention costs in the second block of trials. Alternatively, participants who were free to articulate the stimuli in the first block of trials could have strategically resorted to articulatory simulation, so when suppression was administered in the second block, their performance somewhat suffered. Note that, unlike the *embodiment* view, suppression, in this view, is a strategic adaptation to the experimental setting, not a prerequisite for the computation of syllable structure. Our findings that suppression only emerges in the second block of trials (after the control condition), and it spares sensitivity to syllable structure are in line with this possibility. Together, these results suggest that suppression modulates the identification of speech, but it does not obliterate sensitivity to syllable structure.

Experiment 2: Identity Discrimination

To further investigate the effect of articulatory suppression, we next used the same materials in a second task that potentially imposes greater articulatory demands—identity discrimination. In Experiment 2, participants heard a nonword (the prime, e.g., *blif*) followed (after 3000 ms) by another nonword (the target, e.g., *belif*). Their task was to indicate whether the two items are identical. Unlike Experiment 1, participants in Experiment 2 were required to compare two successive stimuli (e.g., *blif–belif*). To this end, they must maintain the prime *blif* in working memory. And since maintenance in (verbal) working memory requires articulatory rehearsal (e.g., Levy 1971), this task arguably imposes greater articulatory demands compared to the syllable count task (in Experiment 1). Of interest is whether those demands elicit stronger effects of suppression.

If ill-formed onsets are harder to articulate, then monosyllabic primes (e.g., *lbif–lbif*) should elicit stronger articulatory demands than disyllabic primes (e.g., *lebif–lebif*). Consequently, the deleterious effect of suppression should be stronger for monosyllabic primes compared to disyllabic primes. To test this prediction, the following analyses assess the effect of the number of syllables in the prime. Our primary interest is in whether articulatory suppression would attenuate participants' sensitivity to syllable hierarchy.

Method

Participants

A new group of 56 native English speakers (i.e., different from the group who took part in Experiment 1) participated in this experiment. Participants were students at Northeastern University.

Materials

The materials were the same as in Experiment 1, except that they were arranged in pairs. Half of the pairs were physically identical (e.g., monosyllabic: *blif–blif*; disyllabic: *belif–belif*), whereas the other half was nonidentical (e.g., *blif–belif*; *belif–blif*, with order counterbalanced). To counterbalance all the conditions, only 28 quartets[3] (out of 30) were included in this experiment.

Procedure

On each trial, participants heard two auditory nonwords, separated by an SOA of 3000 ms. They were asked to indicate whether the two items are identical by pressing the appropriate key (1 = identical; 2 = nonidentical). Participants were instructed to respond as fast and as accurately as possible.

Results

d Prime Analyses

The effect of suppression on sensitivity (d') was examined by 2 suppression (suppression/control) × 2 prime syllable (the number of syllables in the first stimulus: one/two) × 4 onset type (large rise/small rise/plateau/fall) × 2 list (control−suppression /suppression−control) ANOVAs. The four-way interaction was maringally significant ($F1(3, 162) = 2.96, p < 0.04$; $F2(3, 162) = 2.39, p = 0.071$). Since the list factor is directly linked to block order (i.e., whether the trial occurred in the first or second block of trials; see Table 1), this four-way interaction indicates that suppression modulated responses to the various onset types differently, depending on block order and the number of syllables in the prime word.

To interpret this interaction, we next examined the first and second blocks of trials separately. Of interest is whether suppression modulates performance, and whether suppression attenuates the effect of syllable type. We evaluated these two questions in turn in the following analyses.

a. Does suppression modulate performance? To evaluate the overall effect of suppression, we first compared sensitivity in the suppression and control conditions in the first and second block of trials (collapsed over syllable type). An inspection of the means (Fig. 3) showed that suppression impaired participants' performance. However, this effect was found only in the second block.

We next evaluated the effect of suppression using a 2 suppression × 2 prime syllable × 4 onset type ANOVAs, conducted separately in the two blocks of trials. The effect of suppression in the first block was not significant (both $F < 0.16$), nor did it interact with other factors (for the interactions, all $F < 1.2$). In contrast, results from the second block of trials yielded a significant main effect of suppression ($F1(1, 54) = 2.97, p = 0.09$; $F2(1, 54) = 7.62$, $p < 0.008$), which was not further modulated by onset type or the number of syllables in the prime (for the interactions, all $p > 0.11$).

b. Does suppression attenuate sensitivity to the syllable hierarchy? Given that suppression modulated the identity judgment task, we next examined whether it attenuated sensitivity to

[3] We removed 2 of the quartets (*blif, bnif, bdif, lbif* and *twog, tmok, tpok, mtok*), because the design of this and the following experiments requires that the quartets must be divided evenly across four lists.

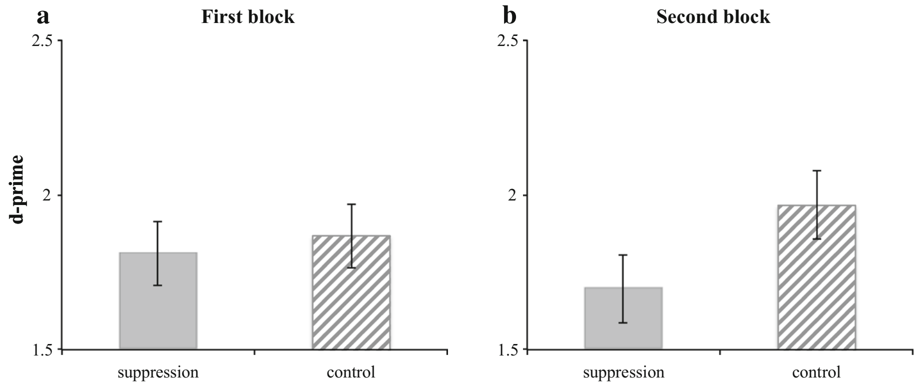


Fig. 3 The effect of articulatory suppression on sensitivity (d') in the first (**a**) and second block (**b**) of Experiment 2. *Note error bars* indicate 95% confidence intervals for the difference between the means

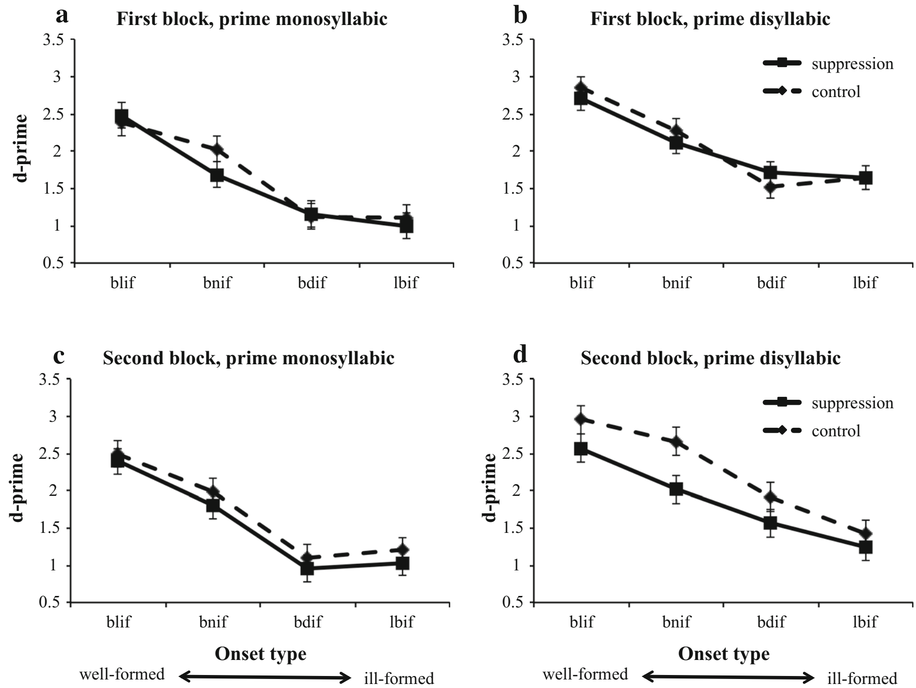


Fig. 4 The effect of articulatory suppression on the sensitivity (d') to onset type in the first (panel **a** and **b**) and the second block (panel **c** and **d**) of Experiment 2. *Note error bars* indicate 95% confidence intervals for the difference between the means

syllable structure. To evaluate this question, we return to the 2 suppression × 2 prime syllable × 4 onset type ANOVAs described in the previous section. Our interest now concerns the effects involving the onset type factor in each of the two blocks of trials.

Analyses of the first block of trials. Figure 4a–b plots the effect of suppression in the first block of trials. The ANOVAs yielded a significant effect of onset type ($F1(3, 162) =$

74.48, $p < 0.0001$; $F2(3, 162) = 44$, $p < 0.0001$), which was not further modulated by either suppression or prime syllable (for the interactions, all $F < 1.2$).

Planned comparisons revealed that sensitivity to *blif*-type syllables was significantly better than to *bnif*-type syllables ($t1(162) = 5.90$, $p < 0.0001$; $t2(162) = 4.55$, $p < 0.0001$). *Bnif*-type onsets, in turn, elicited reliably higher sensitivity than *bdif*-type ones ($t1(162) = 6.63$, $p < 0.0001$; $t2(162) = 5.00$, $p < 0.0001$). Responses to *bdif*- and *lbif*-type syllables did not differ significantly (both $p > 0.72$). Thus, participants were sensitive to most of the syllable hierarchy, and this effect obtained regardless of suppression (i.e., *blif* $\succ$ *bnif* $\succ$ {*bdif*, *lbif*}).

Analyses of the second block. The 2 suppression × 2 prime syllable × 4 onset type ANOVAs yielded a significant a significant main effect of onset type ($F1(3, 162) = 73.78$, $p < 0.0001$; $F2(3, 162) = 60.82$, $p < 0.0001$) as well as a prime syllable × onset type interaction ($F1(3, 162) = 3.48$, $p < 0.02$; $F2(3, 162) = 2.87$, $p < 0.04$). We thus inspected the effect of onset type for monosyllabic and disyllabic primes, separately. The means are presented in Fig. 4c–d.

When prime was *monosyllabic* (see Fig. 4c), onsets with large sonority rises produced reliably better sensitivity than small rises (e.g., *blif* vs. *bnif*, $t1(162) = 3.92$, $p < 0.0002$; $t2(162) = 3.26$, $p < 0.002$). Small rise onsets, in turn, elicited better sensitivity than onsets level in sonority (e.g., *bnif* vs. *bdif*, $t1(162) = 6.15$, $p < 0.0001$; $t2(162) = 5.59$, $p < 0.0001$). Sensitivity to sonority plateaus and falls did not differ significantly (e.g., *bdif* vs. *lbif*, both $p > 0.33$).

Likewise, when prime was *disyllabic*, syllables with large rises (e.g., *blif*) produced significantly higher sensitivity than those with small rises (e.g., *bnif*) ($t1(162) = 3.34$, $p < 0.002$, $t2(162) = 2.92$, $p < 0.005$), which, in turn, elicited significantly better responses than syllables with sonority plateaus (e.g., *bdif*; $t1(162) = 4.71$, $p < 0.0001$, $t2(162) = 4.11$, $p < 0.0001$). Unlike monosyllabic primes, in the case of disyllabic primes, sensitivity to syllables with sonority plateaus (e.g., *bdif*) was also significantly better than those with sonority falls (e.g., *lbif*; $t1(162) = 3.22$, $p < 0.002$; $t2(162) = 2.78$, $p > 0.007$).

Summarizing, the d-prime analyses showed that suppression impaired participants' performance in the discrimination task, in the second block of trials only. Crucially, this effect did not attenuate participants' sensitivity to syllable hierarchy. As the syllables became worse formed, sensitivity declined, regardless of suppression.

Response Time Analyses

We next examined the effects of suppression and onset type on correct response time (RT) to identical (e.g., *blif–blif; belif–belif*) and nonidentical (e.g., *blif–belif; belif–blif*) trials. Unlike the analyses of d', the RT analyses yielded no conclusive evidence for either the effect of suppression or sensitivity to onset structure. The results are detailed below.

Identical Trials. Responses to identical trials were evaluated using 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs (for the means, see Fig. 5a–b and Table 3).

We first asked whether suppression modulated response time. The ANOVAs revealed a significant suppression × list interaction ($F1(1, 54) = 16.0$, $p < 0.0002$; $F2(1, 54) = 17.07$, $p < 0.0002$). However, post hoc tests (Tukey HSD) yielded no reliable differences between the suppression and control conditions in either of the two blocks (all $p > 0.23$). Likewise, none of the effects involving suppression—either main effect (both $F < 1$) or interactions (all $p > 0.32$)—were significant. These findings indicate that suppression did not affect RT to identical trials.

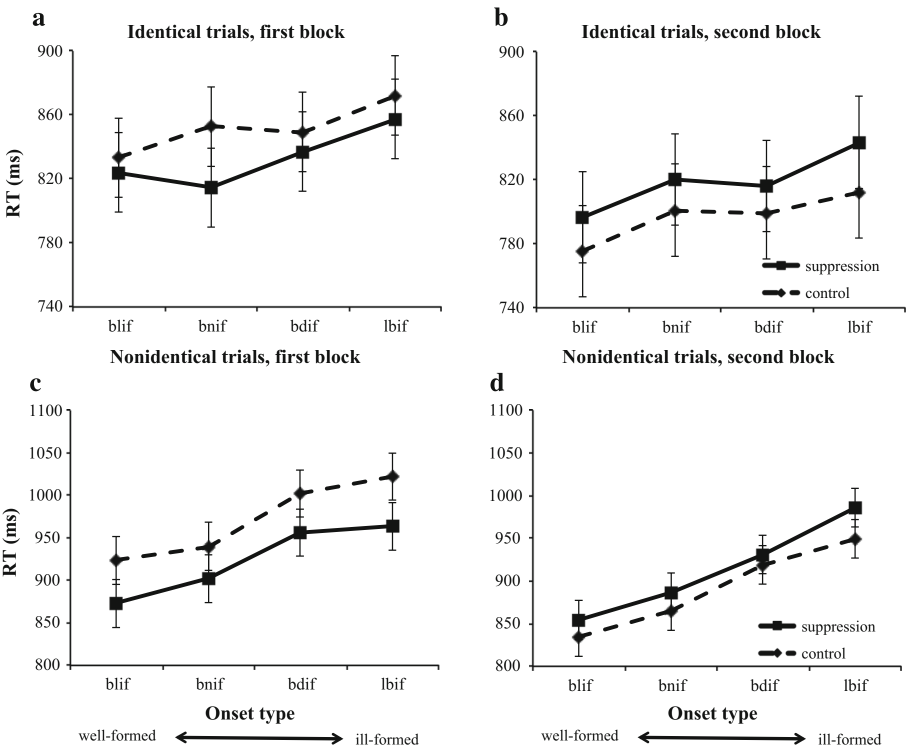


Fig. 5 The effect of suppression on correct response time (RT) to identical (panel **a** and **b**) and nonidentical trials (panel **c** and **d**) of Experiment 2. *Note error bars* indicate 95% confidence intervals for the difference between the means

Moving to the effect of onset type, here, the ANOVAs yielded a significant main effect of onset type ($F1(3, 162) = 13.58$, $p < 0.0001$; $F2(3, 162) = 8.81$, $p < 0.0001$), which was not further modulated by any other factor (all $p > 0.32$). But since the effect of onset type was not specific to monosyllabic items (i.e., there was no prime syllable × onset type interaction), it is not due to the syllable hierarchy *per se*.

Nonidentical Trials. Responses to nonidentical trials (for the means, see Fig. 5c, d, see Table 3) were likewise evaluated by 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs.

Considering first the effect of suppression, the ANOVAs yielded a marginally significant four way interaction ($F1(3, 90) = 1.77$, $p = 0.16$; $F2(3, 93) = 2.72$, $p < 0.05$). In addition, there was a significant suppression × list interaction ($F1(1, 30) = 18.74$, $p < 0.0002$, $F2(1, 31) = 5.27$, $p < 0.03$), suggesting that suppression affected each experimental block differently. Tukey HSD tests showed that responses to the control and suppression manipulation did not differ significantly in either of the blocks (all $p > 0.11$). Similarly, neither the main effect of suppression (both $p > 0.28$) nor its interactions with other factors (all $p > 0.12$) were significant. Thus, suppression did not reliably modulate response time to nonidentical trials.

Moving to the effect of onset type, the ANOVAs yielded a significant main effect of onset type ($F1(3, 90) = 32.75$, $p < 0.0001$; $F2(3, 93) = 29.71$, $p < 0.0001$). Planned compar-

Table 3 Mean response accuracy (ACC, proportion correct) and correct response time (RT) of both identical and nonidentical trials in Experiment 2

	Monosyllabic prime				Disyllabic prime			
	Large rise	Small rise	Plateau	Fall	Large rise	Small rise	Plateau	Fall
ACC								
Identical trials								
First block								
Suppression	0.92	0.84	0.92	0.91	0.89	0.91	0.93	0.99
Control	0.93	0.93	0.94	0.95	0.94	0.95	0.95	0.99
Second block								
Suppression	0.92	0.89	0.90	0.93	0.91	0.93	0.93	0.93
Control	0.93	0.89	0.93	0.97	0.93	0.92	0.97	0.95
Nonidentical trials								
First block								
Suppression	0.82	0.68	0.40	0.37	0.90	0.73	0.56	0.45
Control	0.76	0.66	0.36	0.34	0.88	0.71	0.48	0.45
Second block								
Suppression	0.79	0.64	0.37	0.35	0.84	0.66	0.52	0.43
Control	0.80	0.70	0.37	0.36	0.92	0.86	0.58	0.45
RT (ms)								
Identical trials								
First block								
Suppression	816	812	838	856	832	817	836	859
Control	822	860	849	867	845	845	849	877
Second block								
Suppression	794	819	815	841	799	821	818	845
Control	778	803	795	815	772	799	803	810
Nonidentical trials								
First block								
Suppression	878	909	968	980	866	894	943	946
Control	935	955	1024	1030	911	924	981	1015
Second block								
Suppression	841	903	917	989	867	871	945	983
Control	840	887	933	970	828	841	905	929

isons showed that *blif*-type syllables produced faster responses than *bnif*-type ones ($t1(90) = 1.89$, $p = 0.06$; $t2(93) = 2.66$, $p < 0.01$). *Bnif*-type syllables, in turn, elicited significantly faster responses than *bdif*-type ones ($t1(90) = 3.97$, $p < 0.0002$; $t2(93) = 3.47$, $p < 0.0008$). Finally, *lbif*-type syllables generated reliably slower responses compared to *bdif*-type ones ($t1(90) = 3.27$, $p < 0.0002$; $t2(93) = 2.75$, $p < 0.008$). Thus, as the syllables became worse formed, participants' performance decreased (i.e., *blif* $\succ$ *bnif* $\succ$ *bdif* $\succ$ *lbif*).

Discussion

Experiment 2 extended the investigation of articulatory suppression to an identity discrimination task. The *grammatical* account predicts that participants should be sensitive to onset structure irrespective of suppression. The *motor embodiment* account asserts otherwise—suppression should decrease *overall* performance but it should *improve* the identification of ill-formed syllables, especially when those syllables are presented first (as the prime). Consequently, participants' sensitivity to syllable hierarchy should be attenuated.

Our results showed that suppression indeed *impaired* performance in the second block of trials. When suppression was administered, responses were reliably slower and more error-prone. Contrary to the *motor embodiment* account, however, the effect of suppression was not larger for monosyllabic primes. Furthermore, rather than improving performance with ill-formed syllables (by releasing ill-formed syllables like *lbif* from the costs of articulatory simulation), suppression *impaired* the discrimination of better-formed ones (e.g., *blif* and *bnif*). Crucially, suppression did not attenuate participants' sensitivity to onset structure. So while participants were sensitive to suppression (possibly, due to the conjunction of its attention demands, fatigue and strategic reliance on articulation, as detailed in the Discussion of Experiment 1), these findings are inconsistent with the *motor embodiment* account.

Experiment 3: Identity Discrimination with Printed Materials

Results from Experiments 1–2 suggest that the effect of the syllable hierarchy is inexplicable by articulatory factors. However, both experiments used auditory materials. Accordingly, one might worry that the effect of syllable structure reflects difficulties in extracting auditory/phonetic representations. For example, people might misidentify *lbif* because its phonetic form is confusable with *lebif*. The misidentification of ill-formed auditory syllables could thus be unrelated to grammatical restrictions.

To address this possibility, past research has examined the syllable hierarchy using printed materials. This work builds on the ample reading literature, showing that readers assemble phonological representations from print in silent reading (e.g., Berent and Perfetti 1995; Van Orden et al. 1990). Printed words thus allow one to assess the effect of phonological constraints while controlling the phonetic demands of auditory stimuli. If readers are sensitive to the syllable structure of printed materials, then this effect is unlikely due to auditory/phonetic failure. Results indeed showed that, as the syllable became worse formed on the syllable hierarchy, discrimination difficulties increased (Berent 2008; Berent et al. 2009, 2014). These findings clearly counter the auditory/phonetic explanation. Nonetheless, these effects of syllable structure could emanate from articulatory simulation. Indeed, past research has demonstrated that subvocal articulation indeed plays an important role in silent reading (e.g., Eiter and Inhoff 2008; Lukatela et al. 2004; Besner 1987; Baddeley et al. 1975; Besner and Davelaar 1982; Hitch and Baddeley 1976; Kleiman 1975; Levy 1975, 1978; Martin 1978). In view of these findings, it is possible that readers' sensitivity to the syllable hierarchy reflects articulatory simulation.

To further investigate the source of the syllable hierarchy, we next asked whether sensitivity to syllable structure is maintained even when participants identify printed materials under articulatory suppression. The *grammatical* (but not the *motor embodiment*) account predicts that sensitivity to the syllable hierarchy should be maintained irrespective of suppression. However, the manifestation of this effect with printed words might be more complex than with auditory materials.

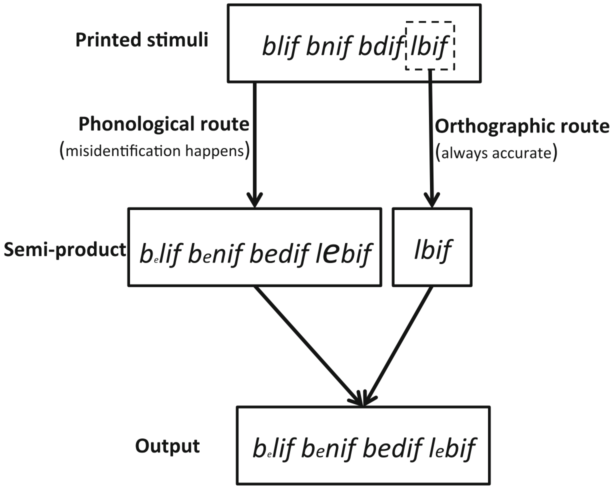


Fig. 6 A dual-route account for the identification of printed materials in Experiments 3–4

Unlike auditory stimuli, printed monosyllables could be discriminated from their disyllabic counterparts using multiple sources of information—either through phonological decoding or an orthographic route (Fig. 6). Furthermore, participants could strategically shift their reliance on the two routes depending on their overall reliability. Indeed, orthographic information can reliably distinguish monosyllables from disyllables. In contrast, phonological decoding is an elaborate process that is error-prone—the worse formed a monosyllable, the more likely its misidentification as a disyllable (e.g., *lbif* $\rightarrow$ *lebif*). To avoid phonological errors, participants might resort to graphemic verification strategies. For example, to determine whether *lbif* is identical to *lebif*, participants could simply monitor the letter *e* in the second letter-position. The shift from phonological decoding to orthographic strategies is especially likely for sonority falls, as their phonological decoding is the most error-prone, and their graphemic structure is distinct (i.e., they begin with a sonorant consonant). If sonority falls elicit a shift from the (error-prone) phonological decoding to the (accurate) graphemic verification, then the effect of the syllable hierarchy on performance may be nonlinear—an effect documented in past research (Lennertz and Berent 2015). Sonority falls should yield relatively accurate responses (via the graphemic route). By contrast, the identification accuracy of better-formed syllables should be mediated by phonology assembly, hence, accuracy should decline as they become worse formed (i.e., *blif* $\succ$ *bnif* $\succ$ *bdif*). Experiment 3 tested these predictions.

Method

Participants

A new group of 48 native English speakers, students of Northeastern University, participated in this experiment.

Materials

A printed version of the same nonword stimuli from the previous experiments was included, arranged as explained in Experiment 2. To minimize the effect of visual overlap, prime and target were presented in different type cases, masked by a series of Xs.

Procedure

Participants initiated each trial by pressing the space bar. Their response triggered the presentation of nonword prime, displayed for for 500 ms, followed by a mask of "XXXXXXX", displayed for 2500 ms, and finally, the nonword target, presented for up to 2500 ms. Participants were instructed to quickly indicate whether the two nonwords (the prime and the target) were identical by pressing one of two keys (1 = identical, 2 = nonidentical). Note that the SOA in this experiment (3000 ms) matched the SOA used in Experiment 2.

Results

d-Prime Analyses

An inspection of the means (Fig. 7) suggests that suppression diminished overall sensitivity (d'), but participants were still sensitive to onset structure even under suppression. As expected, however, the effect of onset type was nonlinear. Specifically, when presented with obstruent-initial syllables (e.g., *blif,bnif,bdif*), sensitivity in the suppression condition gradually declined as the syllable became worse formed along the syllable hierarchy. However, the worst-formed *lbif*-type syllables was associated with better discrimination.

The 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs yielded a significant main effect of suppression ($F1(1, 46) = 12.83, p < 0.0009$; $F2(1, 54) = 22.75, p < 0.0001$). These findings indicate that the suppression manipulation impaired participants' overall ability to discriminate monosyllables from their disyllabic counterparts.

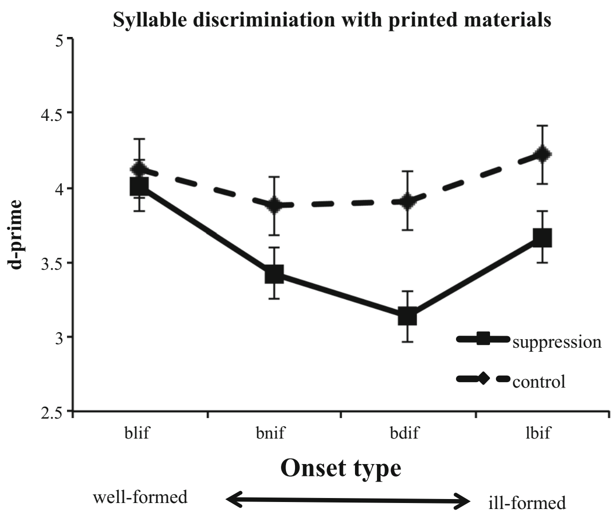


Fig. 7 The effect of suppression and onset type on sensitivity (d') in Experiment 3. *Note error bars* indicate 95% confidence intervals for the difference between the means

The ANOVAs also yielded a reliable effect of onset type ($F1(3, 138) = 7.73$, $p < 0.0001$; $F2(3, 162) = 7.77$, $p < 0.0001$). The onset type × list interaction was significant only across participants, $F1(3, 138) = 3.08$, $p < 0.03$; but not by item, $F1(3, 162) = 2.00$, $p > 0.11$), and it was not further modulated by suppression (for the suppression × onset type interaction, $F1(3, 138) = 2.23$, $p = 0.087$; $F2(3, 162) = 1.95$, $p > 0.12$) or any other factor (all $p > 0.12$).

Planned comparisons showed that onsets with large sonority rises (e.g., *blif*) elicited significantly better sensitivity than those with small rises (e.g., *bnif*; $t1(138) = 3.26$, $p < 0.002$; $t2(162) = 3.03$, $p < 0.003$). Likewise, onset with small rises (e.g., *bnif*) elicited numerically better sensitivity than those with sonority plateaus (e.g., *bdif*; both $p > 0.32$, *n.s.*). Notably, onsets with sonority plateaus (e.g., *bdif*) elicited significantly *worse* performance compared to sonority falls (e.g., *lbif*; $t1(138) = 3.27$, $p < 0.002$; $t2(162) = 3.67$, $p < 0.0004$).

Thus, the suppression manipulation effectively impaired readers' ability to distinguish monosyllables from disyllables. However, participants remained sensitive to syllable hierarchy even when suppression was administered: as the onset became worse formed, discrimination generally decreased. The one notable exception was presented by the benefit of *lbif*-type onsets.

Response Time Analyses

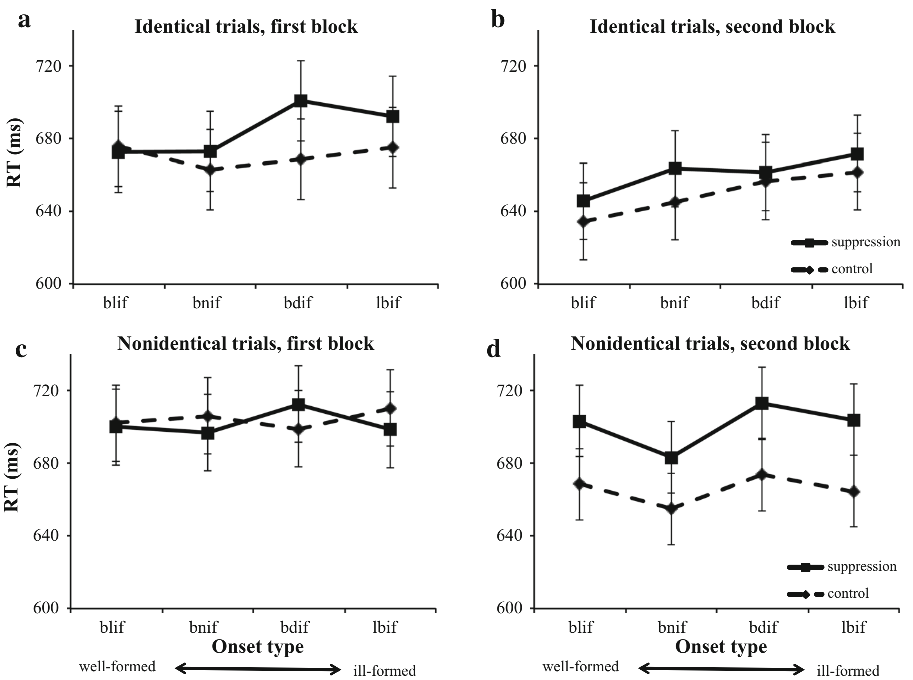
We next evaluated the effects of suppression and onset type on correct response time (RT) to identical and nonidentical trials. Unlike the findings from the d' analyses, the RT data did not reveal reliable effects of suppression or onset type. These analyses are described below; the results are plotted in Fig. 8.

Identical Trials. The 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs on accurate responses to identical trials yielded a significant suppression × list interaction ($F1(1, 45) = 6.41$, $p < 0.02$; $F2(1, 54) = 4.18$, $p < 0.05$). However, Tukey HSD tests showed that the control vs. suppression contrasts were not reliable in either block[4] (all $p > 0.79$). In addition, there was no main effect of suppression, nor did suppression interact with any other factors (all $p > 0.1$). These findings indicate that suppression did not affect RT to identical trials.

The ANOVAs also yielded a significant main effect of onset type ($F1(3, 135) = 3.63$, $p < 0.02$; $F2(3, 162) = 3.67$, $p < 0.02$), and it was not further modulated by prime syllable ($F1(3, 135) = 1.58$, $p > 0.19$; $F2(3, 162) = 2.66$, $p = 0.05$) or any other factors (for all interactions, $p > 0.16$). Planned comparisons, however, revealed no reliable contrasts between any of the adjacent onset types.

Nonidentical Trials. The 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs produced a significant suppression × list interaction ($F1(1, 46) = 4.71$, $p < 0.04$; $F2(1, 54) = 4.53$, $p < 0.04$). Post-hoc tests (Tukey HSD) showed that the control vs. suppression contrast was marginally significant in the second block (by item, $p < 0.02$; albeit not by participant, $p > 0.5$)—responses were significantly faster under the control than the suppression condition. The control-suppression contrast was not significant in the first block (both $p > 0.99$). In addition, there was no reliable effect of suppression

[4] Block order (i.e., whether the trial appeared in the first or second block of trials) is captured by the list factor (see Table 1 for details).



Discussion

Fig. 8 The effect of suppression and onset type on correct response time (RT) to identical (*panel* **a** and **b**) and nonidentical trials (*panel* **c** and **d**) in Experiment 3. *Note error bars* indicate 95% confidence intervals for the difference between the means

($F1(1, 46) = 3.24$, $p = 0.08$; $F2(1, 54) = 2.68$, $p > 0.1$), and it did not interact with any other factors (for all interactions with suppression, $p > 0.22$).

Unlike the identical trials, however, there was no effect of onset type in responses to the nonidentical trials (both $p > 0.2$), nor was there an interaction between onset type and other factors (all $p > 0.11$).

Discussion

Experiment 3 gauged the effect of syllable structure with printed materials. Results showed that suppression indeed elicited a *decrease* in readers' ability to discriminate monosyllables from their disyllabic counterparts, an observation in line with *motor embodiment* account. Nonetheless, when considering the better-formed (i.e., obstruent-initial) syllables, readers were sensitive to the syllable hierarchy regardless of suppression (i.e., $blif \succ bnif \succ bdif$).[5] Notably, the worst-formed syllables like *lbif* produced *higher* sensitivity than the better-formed *bdif*-type ones (i.e., $lbif \succ bdif$, rather than $bdif \succ lbif$), and this advantage obtained

[5] To further ascertain readers' sensitivity to onset structure under suppression, we next tested for the effect of syllable type under the suppression condition, separately. Planned comparisons showed that *blif*-type syllables yielded significantly better sensitivity than *bnif*-type ones ($t1(138) = 3.20$, $p < 0.002$; $t2(162) = 3.01$, $p < 0.003$). Sensitivity to *bnif*-type syllables, in turn, was numerically higher than *bdif*-type onsets (both $p > 0.10$, *n.s.*). This aspect of our findings is inconsistent with the *motor embodiment* account.

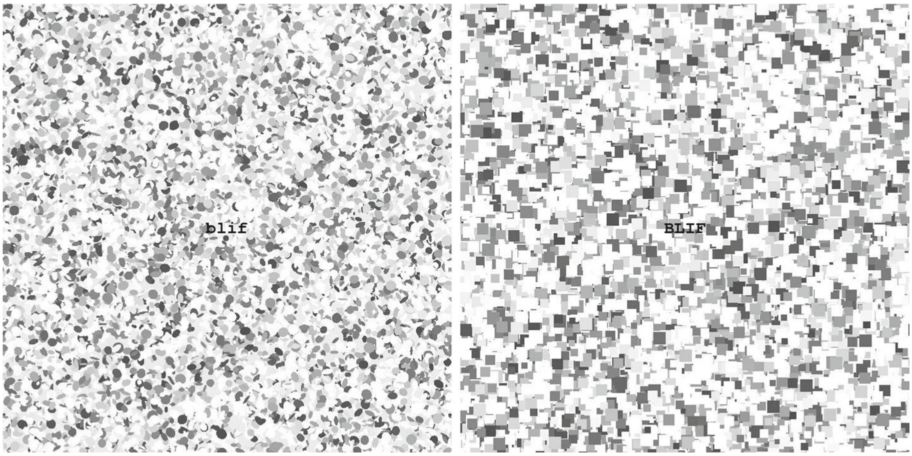


Fig. 9 An illustration of the stimulus presentation in Experiment 4 (*left* prime blif; *right* target BLIF)

irrespective of articulatory suppression. We suggest that this reversed effect of onset type is due to the contribution of orthographic information.

As discussed earlier, printed materials afford two possible routes for syllable identification—either through phonological decoding or by spelling verification. The phonological decoding process is error-prone, whereas spelling verification is highly accurate. If participants are aware of their tendency to misidentify sonorant-initial clusters (e.g., *lbif*), then an encounter with such syllables might prompt them to rely on a direct orthographic process. And because spelling provides unambiguous cues to the number of syllables, responses to ill-formed syllables should become more accurate. The facilitation for the worst-formed syllables like *lbif* could thus reflect a strategic adjustment, adopted as a direct consequence of their ill-formedness. Experiment 4 further tests this possibility.

Experiment 4: Identity Discrimination with Printed Materials and Background Noise

Experiment 4 examines whether the advantage of *lbif*-type syllables is due to the adoption of a graphemic verification strategy. To this end, we presented the materials (the printed stimuli from Experiment 3) occluded by visual noise (see Fig. 9). In so doing, we wished to discourage participants from relying on graphemic verification. If the advantage of *lbif*-type syllables reflected a graphemic strategy, then this advantage should be attenuated by visual noise. Consequently, these items should now produce lower sensitivity relative to better-formed items (as predicted by the *grammatical* account).

Method

Participants

Another 32 Northeastern University undergraduate students took part in this experiment. All of them were native English speakers.

Materials

The materials of this experiment consisted of nonword stimuli presented against black-and-white background visual noise. The nonword stimuli were those used in Experiment 3, and the discrimination task is otherwise identical to the one used in Experiment 3. The background noise images were a collection of randomly distributed discrete objects that varied in size and contrast. All of these images were generated by a dead leaves model (Lee et al. 2001). To better obscure the shape of the letters, two distinct types of patterns were included—either filled with circular- or square-shaped objects (see Fig. 9 left and right panel, respectively). As most of the lower- and upper-cased letters of our materials were composed of circular and angular shapes, respectively, the lower-case stimuli (i.e., the primes; e.g., blif) were presented with images filled with round objects, whereas the upper-case stimuli (i.e., the targets; e.g., BELIF) were displayed with patterns of squares. Each circular pattern was randomly paired with one square pattern into a pair. A total of 32 such pairs were included. We randomly assigned 4 pairs to the practice trials and the other 28 to the experimental trials. Each image pair was randomly selected to present with one nonword quartet.

Procedure

The procedure was exactly the same as that of Experiment 3.

Results

Our analyses proceed in two steps. We first examined whether visual noise decreases or eliminates the advantage of worst-formed syllables like *lbif*. Inasmuch as the background interference is effective, we next asked whether readers remain sensitive to the syllable hierarchy despite articulatory suppression.

1. Are Readers Sensitive to Visual Noise?

To determine whether our background noise manipulation was effective, we first compared readers' sensitivity (d') in Experiments 3 and 4 via 2 experiment $\times$ 2 suppression $\times$ 2 prime syllable $\times$ 4 onset type $\times$ 2 list ANOVAs. The significant main effect of experiment ($F1(1, 76) = 89.15$, $p < 0.0001$; $F2(1, 108) = 405.33$, $p < 0.0001$), showed that performance was impaired by visual noise (in Experiment 4 relative to the no noise in Experiment 3).

The ANOVAs also yielded a marginally significant interaction of suppression $\times$ onset type $\times$ experiment $\times$ list interaction ($F1(3, 228) = 2.59$, $p = 0.054$; $F2(3, 324) = 2.14$, $p = 0.095$). The 5-way interaction (suppression $\times$ prime syllable $\times$ onset type $\times$ experiment $\times$ list), was significant by items ($F1 < 1$; $F2(3, 324) = 2.58$, $p = 0.05$); the means are plotted in Fig. 10.

Recall, however, that in Experiment 3, the interaction between suppression and onset type was not modulated by the experimental list. Accordingly, the high order interaction across experiments (suppression $\times$ onset type $\times$ experiment $\times$ list) must specifically come from Experiment 4, possibly due to the addition of visual noise. To further investigate the effect of visual noise, we examined results from Experiment 4 separately. Of interest is whether noise eliminated the anomalous advantage of the worst formed syllables (e.g., *lbif*).

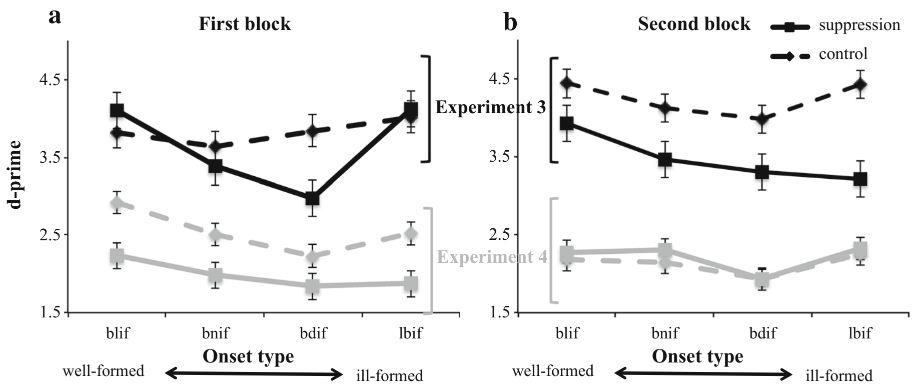


Fig. 10 The effect of articulatory suppression and onset type on sensitivity (d') under the noise-absent and the noise-present conditions (Experiment 3 and 4, respectively). *Note error bars* indicate 95% confidence intervals for the difference between the means

2. Does Suppression Attenuate Sensitivity to the Syllable Hierarchy?

In this section, we examined the effect of suppression and onset type under visual noise, in Experiment 4. Two sets of analyses were conducted, one on d-prime and the other on the RT data.

An inspection of the d-prime means (Fig. 11) suggests that suppression impaired syllable discrimination, but this effect only emerged in the first block of trials. As in Experiment 3, readers remained sensitive to most of the syllable hierarchy regardless of suppression. While the worst-formed onsets (e.g., *lb*) improved discrimination in the second block of trials, this effect was eliminated when suppression was administered in the first block. Our subsequent analyses tested the effect of suppression and the sensitivity to syllable structure.

The 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs yielded a significant main effect of suppression ($F1(1, 30) = 5.53$, $p < 0.03$; $F2(1, 54) = 10.59$, $p < 0.002$), suggesting the suppression manipulation impaired discrimination. In addition, the effect of suppression was modulated by the number of syllables in the prime (for suppression × prime syllable interaction, $F1(1, 30) = 10.26$, $p < 0.004$; $F2(1, 54) = 11.14$, $p < 0.002$) and further by experimental list (for the suppression × prime syllable × list interaction, $F1(1, 30) = 5.21$, $p < 0.03$; $F2(1, 54) = 5.05$, $p < 0.03$)[6]. The 4-way suppression × prime syllable × onset type × list interaction, however, was not significant ($F1(3, 90) = 1.42$, $p = 0.24$; $F2(3, 162) = 2.31$, $p = 0.079$). An inspection of the means suggests that

[6] To further examine the effect of suppression, we further probed the suppression × prime syllable × list interaction, by analyzing the two experimental blocks separately. An inspection of the means showed that the effect of suppression was limited to monosyllabic primes in the first block.
In the first block of trials, the 2 suppression × 2 prime syllable × 4 onset type ANOVAs produced a significant main effect of suppression ($F1(1, 30) = 8.87$, $p < 0.006$; $F2(1, 54) = 21.19$, $p < 0.0001$) and a reliable interaction between suppression and prime syllable ($F1(1, 30) = 5.65$, $p < 0.03$; $F2(1, 54) = 16.08$, $p < 0.0002$). Planned comparisons showed that compared to control condition, suppression significantly decreased discrimination, but only when the prime was monosyllabic ($t1(50) = 3.78$, $p < 0.0005$; $t2(88) = 6.01$, $p < 0.0001$). When prime was disyllabic, responses to the control and suppression conditions did not differ significantly (the contrast was only marginally significant by items, $t2(88) = 1.9$, $p < 0.07$; by participants, $p = 0.21$, $n.s.$). In the second block, the 2 suppression × 2 prime syllable × 4 onset type ANOVAs showed no effect of suppression. There was no main effect of suppression (both $F < 1$), nor was it modulated by prime syllable, or onset type (for interactions, all $p > 0.21$).

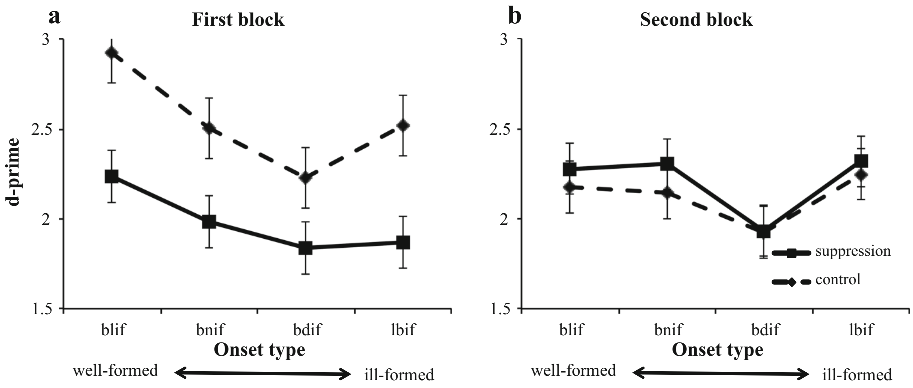


Fig. 11 The effect of suppression and onset type on sensitivity (d') in the first and the second block (*panel* **a** and **b**, respectively) of Experiment 4. *Note error bars* indicate 95% confidence intervals for the difference between the means

the effect of suppression was limited to monosyllabic primes in the first block. Additional follow-up test of this interaction are described in footnote 6.

Given that suppression modulated performance in this task, we next asked whether suppression further attenuated the effect of syllable structure. The ANOVAs yielded a reliable suppression × onset type × list interaction ($F1(3, 90) = 2.70$, $p = 0.05$; $F2(3, 162) = 2.13$, $p = 0.1$). Since the list factor captures block order (see Table 1), we next investigated the effect of suppression and onset type in the first and second blocks separately via 2 suppression × 2 prime syllable × 4 onset type ANOVAs (for the means, see Fig. 11a–b and Table 4).

In the first block of trials, the ANOVAs yielded a significant main effect of onset type ($F1(3, 90) = 6.05$, $p < 0.001$; $F2(3, 162) = 5.19$, $p < 0.002$), that did not interact with other factors (all $p > 0.46$). Planned comparisons showed that *blif*-type syllables elicited significantly better sensitivity than *bnif*-type ones ($t1(90) = 2.55$, $p < 0.02$; $t2(162) = 2.60$, $p < 0.02$) and *bdif*-type items ($t1(90) = 4.14$, $p < 0.0001$; $t2(162) = 3.70$, $p < 0.001$). Other comparisons (e.g., *bnif vs. bdif* and *bdif vs. lbif*), however, were not statistically significant (all $p > 0.11$). These results demonstrate that in the first block of trials, people were sensitive to onset structure. Furthermore, the anomalous advantage of *lbif*-type syllables was eliminated.

Similar analyses conducted in the second block of trials also yielded a marginally significant main effect of onset type ($F1(3, 90) = 3.60$, $p < 0.02$; $F2(3,162) = 2.14$, $p = 0.097$), which was not modulated by any other factors (for all interactions, $p > 0.28$). Planned comparisons revealed that the discrimination of *blif*-type onsets did not differ from *bnif*-type items ($p > 0.8$), which, in turn, elicited significantly better sensitivity than *bdif*-type ones ($t1(90) = 2.47$, $p < 0.016$; $t2(162) = 1.76$, $p = 0.08$). Unlike the first block of trials however, in the second block, the worst-formed *lbif*-type syllables produced significantly *better* discrimination accuracy relative to *bdif*-type ones ($t1(90) = 2.97$, $p < 0.004$; $t2(162) = 2.31$, $p < 0.03$)—a result that mirrors the findings from Experiment 3.

Additional analyses on RT[7] did not yield any effects of suppression or onset type. The RT means are provided in Table 4.

[7] The effect of suppression and onset type was examined for identical and nonidentical trials separately. For identical trials, the 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVAs yielded a significant suppression × list interaction ($F1(1, 30) = 7.06$, $p < 0.013$; $F2(1, 54) = 10.90$, $p < 0.002$). However, post-hoc comparisons found no significant effect of suppression in either block (all $p > 0.35$). Likewise, none

Table 4 Mean response accuracy (ACC, proportion correct) and correct response time (RT) of nonidentical trials in Experiment 4

	Monosyllabic prime				Disyllabic prime			
	Large rise	Small rise	Plateau	Fall	Large rise	Small rise	Plateau	Fall
ACC								
Identical trials								
First block								
Suppression	0.88	0.88	0.81	0.77	0.88	0.88	0.83	0.82
Control	0.94	0.96	0.92	0.88	0.89	0.88	0.87	0.89
Second block								
Suppression	0.86	0.86	0.87	0.89	0.83	0.85	0.80	0.82
Control	0.84	0.83	0.85	0.88	0.84	0.84	0.80	0.81
Nonidentical trials								
First block								
Suppression	0.76	0.63	0.69	0.71	0.83	0.80	0.81	0.85
Control	0.92	0.82	0.75	0.85	0.95	0.83	0.82	0.87
Second block								
Suppression	0.80	0.82	0.73	0.84	0.88	0.87	0.81	0.86
Control	0.79	0.83	0.78	0.86	0.85	0.79	0.76	0.80
RT (ms)								
Identical trials								
First block								
Suppression	735	713	728	724	728	727	739	752
Control	709	682	700	744	706	705	747	736
Second block								
Suppression	670	720	652	721	726	693	728	736
Control	668	688	683	670	694	705	679	699
Nonidentical trials								
First block								
Suppression	796	788	802	824	772	777	753	804
Control	761	800	776	819	758	741	743	744
Second block								
Suppression	746	716	757	740	745	740	747	718
Control	732	748	746	751	773	721	737	764

In summary, our results suggest two conclusions. First, the visual noise manipulation effectively decreased readers' discrimination accuracy, and with it, the anomalous *lbif*-

Footnote 7 continued
of the effects (main effect or interactions) involving the onset type factor were significant (all $p > 0.07$;). In the nonidentical trials, the 2 suppression × 2 prime syllable × 4 onset type × 2 list ANOVA also exhibited a significant suppression × list interaction ($F1(1, 30) = 10.42$, $p < 0.004$; $F2(1, 54) = 7.86$, $p < 0.008$), and it was further modulated by prime syllable (suppression × prime syllable × list interaction, $F1(1, 30) = 5.86$, $p < 0.03$; $F2(1, 54) = 4.33$, $p < 0.05$). We probed the 3-way interaction by investigating each block separately. The effect of suppression was not significant in either block (all $p > 0.16$), and it was not reliably modulated by other factors (all $p > 0.23$). Furthermore, the main effect of onset type was not significant in either block (all $p > 0.11$), nor did it interact with any other factors (all $p > 0.21$).

advantage was completely eliminated in the first block of trials. Second, readers were sensitive to most of the syllable hierarchy regardless of suppression. Specifically, as the syllables became better formed, discrimination of obstruent-initial clusters (e.g., *bl, bn, bd*) improved.

Discussion

Findings from Experiment 3 indicate that *lbif*-type syllables elicited unexpectedly better discrimination relative to *bdif*-type ones. We suggest this benefit is due to a graphemic verification strategy, triggered by the phonological ill-formedness of such onsets. To test this explanation, Experiment 4 discouraged the use of this strategy by obscuring the printed items with noisy background patterns. If the anomalous advantage of *lbif*-type syllables in Experiment 3 is due to a visual strategy, then this advantage should be attenuated by visual noise, and their expected grammatical dispreference relative to better-formed syllables should now emerge.

Comparing Experiments 3 and 4, we found that visual noise decreased discrimination accuracy. Critically, the noise manipulation reduced the advantage of *lbif*-type syllables in discrimination—this advantage was completely eliminated in the first block of trials and it only emerged in the second block. Such late emergence could have occurred because visual verification is a controlled (i.e., attention demanding) strategy that develops over time. Crucially, in the absence of learning (i.e., in the first block), and under the suppression load, the *lbif* advantage disappeared.

We should note, however, that the *grammatical* account predicts that *lbif*-type items should result in the lowest discrimination accuracy. The results of Experiment 4 did not support this prediction. It is unclear whether this finding reflects residual effects of graphemic verification under visual noise, or some other factors—further research is necessary to evaluate this issue. Clearly, however, readers remained sensitive to the syllable hierarchy despite articulatory suppression. This aspect of our findings challenges the *motor embodiment* account.

General Discussion

Across languages, certain syllable types are systematically preferred to others (e.g., *blif* $\succ$ *bnif* $\succ$ *bdif* $\succ$ *lbif*). Similar preferences are evident in the behavior of individual speakers (English: Berent et al. 2007; Spanish: Berent et al. 2012; French: Maïonchi-Pino et al. 2012; Hebrew: Berent et al. 2013). Results show that, as the syllable becomes worse formed (i.e., dispreferred), misidentification rate systematically increases.

The misidentification of ill-formed syllables is inexplicable by lexical analogy (e.g., *bnif* is preferred because it is similar to *sniff*), as similar findings emerge even in languages that lack onset clusters altogether (Korean: Berent 2008; Mandarin: Zhao and Berent 2015). It is also unlikely that these preferences are solely due to auditory/phonetic reasons, as sensitivity to syllable struture obtained with printed materials (Berent 2008; Berent and Lennertz 2010; Tamási and Berent 2014). However, the existing literature leaves open the possibility that misidentification is due to motor demands imposed by articulatory simulation (i.e., *motor embodiment*). To evaluate this possibility, the present research investigated whether people remain sensitive to onset structure despite articulatory suppression.

If the syllable hierarchy results from grammatical constraints, as predicted by the *grammatical* account, then our participants should respect this hierarchy even when articulatory activity is suppressed. By contrast, if the syllable hierarchy arises from motor demands, as asserted by the *motor embodiment* account, then sensitivity to this hierarchy should be attenuated under articulatory suppression.

In line with the *motor embodiment* hypothesis, results from all four experiments showed that articulatory suppression indeed impaired participants' overall performance. Contrary to this hypothesis, however, participants remained highly sensitive to the syllable structure of auditory stimuli (Experiments 1 and 2) irrespective of suppression. Moreover, the effect of suppression with auditory stimuli emerged only in the second block of trials. These findings suggest that during syllable identification/discrimination, listeners do not engage in articulatory simulation automatically; rather, simulation is a controlled process that develops throughout the experimental session. Motor simulation, therefore, is unlikely the sole source of syllable hierarchy.

It is still possible, however, that the findings with auditory materials arise from difficulties to extract auditory/phonetic cues. To address this possibility, we extended our investigation to printed words. When presented visual stimuli (Experiments 3 and 4), our findings were more complex. In line with the *grammatical* account, in each experiment, readers respected part of the syllable hierarchy (i.e., $bl \succ bn \succ bd$), and this effect was found regardless of suppression. This aspect of the findings is clearly inconsistent with either the auditory/phonetic or the *motor embodiment* explanations. Surprisingly, the worst formed syllables (e.g., *lbif*) produced the most accurate discrimination.

We suggest that this unexpected advantage is due to a visual strategy of spelling verification. Because spelling offers reliable cues for discriminating monosyllables (e.g., *lbif*) from their disyllabic counterparts (e.g., *lebif*), this strategy should yield relatively accurate responses. Worst-formed syllables (e.g., *lbif*) are most likely to invoke this strategy since their phonological decoding is error-prone, on the one hand, and since their unique graphemic/phonological structure (i.e., sonorant-initial clusters) distinguishes them from all other monosyllables, on the other. The selective spelling verification should thus confer an advantage to *lbif*-type syllables, and this effect could have masked their grammatical ill-formedness. In line with this explanation, we found that the anomalous advantage of *lbif*-type stimuli was eliminated in the presence of visual noise (in the first block of Experiment 4). Together, these results suggest that the syllable hierarchy reflects neither articulatory demands nor auditory/phonetic difficulties in processing auditory materials, but rather to their grammatical phonological structure.

To further evaluate the grammatical explanation, we next contrast the grammatical account with two non-grammatical alternatives via additional regression analyses of our findings. One set of analyses examines whether the effect of syllable type with auditory stimuli (in Experiments 1–2) is due to difficulty to encode their phonetic cues; a second set of analyses examines whether the advantage of better-formed syllables across our four experiments is due to lexical analogy.

1. Is the Misidentification of Ill-Formed Auditory Syllables Due to Auditory/Phonetic Failure?

To further test the possibility that the misidentification of ill-formed syllables is due to difficulties to encode their auditory/phonetic form, we submitted our findings with the auditory stimuli (Experiments 1 and 2) to a series of step-wise regression analyses.

These analyses gauge the effect of the duration and intensity of the burst release associated with stop consonants (e.g., *b* in *blif*). Past research (Kang 2003; Wilson and Davidson 2013; Wilson et al. 2014) has shown that participants sometimes misinterpret the burst as evidence for an intermediate schwa, and consequently, they misidentify the monosyllabic input as disyllabic (e.g., as *belif*). To evaluate this phonetic explanation, we next compared the unique phonetic effect of the burst (its intensity and duration) with the effect of onset type. The definition and measures of burst were obtained from Berent et al. (2008), and the summary statistics of these measures are provided in Table 5.

Our analyses examined the effect of the ***phonetic properties*** of the burst on sensitivity (d') in the suppression and control conditions, separately. To test participants' sensitivity to phonetic cues, we first entered onset type in the first step, and then forced burst duration and intensity (together) as the last predictor. Results (Table 6) showed that, in the control condition, phonetic cues uniquely captured less than 8% of the variance (Experiment 1: $\Delta R^2 = 0.032$; Experiment 2: $\Delta R^2 = 0.079$), and this effect was either significant (in Experiment 2) or marginally significant (in Experiment 1). Interestingly, once suppression was administered, the unique effect of phonetic cues was no longer reliable. Thus, participants in our experiments were sensitive to phonetic cues, but this effect was eliminated (rather than increased) by articulatory suppression.

A second set of analyses next investigated whether participants were also sensitive to the ***phonological structure*** of the syllable. To assess the unique phonological effect of onset type, we reversed the order of predictors. In these models, the phonetic properties of the burst were entered first, whereas onset type was entered last. We found that participants were highly sensitive to onset type even after controlling for phonetic factors, and this effect obtained in both experiments irrespective of suppression. Not only did the unique effect of onset type survive suppression, but its size (all $\Delta R^2 > 0.46$) was far larger than that of the burst (all $\Delta R^2 < 0.08$).

We conclude that the phonological effect of onset type is not subsumed by the phonetic properties of the burst, and it obtains irrespective of articulatory suppression. The insufficiency of auditory/phonetic factors to capture the syllable hierarchy in the present experiments is consistent with past research with printed materials (e.g., Lennertz 2010, Experiment 4; Tamási and Berent 2014, Experiment 3). These results showed that people respect the full syllable hierarchy even in the absence of auditory input, and even when lexical factors are statistically controlled (via regression analyses). Additional challenges to the auditory/phonetic account are presented by the documentation of the syllable hierarchy among people with dyslexia—individuals whose auditory/phonetic systems are demonstrably impaired (Berent et al. 2013; Berent et al. 2016). Together, these findings suggest that syllable hierarchy reflects abstract linguistic restrictions, rather than phonetic/auditory difficulties alone.

Table 5 Duration (ms) and intensity (dB) of burst release in Experiments 1–2

Experiment	Onset type	Burst duration (ms)		Burst intensity (dB)	
		Mean	SD	Mean	SD
1	blif	10.11	3.14	60.5	5.42
	bnif	9.58	3.29	59.07	6.29
	bdif	10.28	4.67	62.04	5.03
2	blif	9.92	3.15	60.4	5.42
	bnif	9.53	3.39	59.09	6.4
	bdif	9.61	3.43	61.65	4.97

Table 6 The unique effect of (A) phonetic cues (burst intensity and duration); and (B) onset type in step-wise regression analyses of sensitivity (d') in Experiments 1–2

Experiment	Condition	Step	Predictor	ΔR^2	ΔF	*df*	*P* value
1	Control	1	Onset type	0.517	94.078	1.88	0***
		2	Burst intensity and duration	0.032	3.076	2.86	0.051†
		1	Burst intensity and duration	0.059	2.726	2.87	0.071†
		2	Onset type	0.49	93.425	1.86	0***
	Suppression	1	Onset type	0.592	127.75	1.88	0***
		2	Burst intensity and duration	0.023	2.539	2.86	0.085†
		1	Burst intensity and duration	0.054	2.469	2.87	0.091†
		2	Onset type	0.561	125.306	1.86	0***
2	Control	1	Onset type	0.488	78.222	1.82	0***
		2	Burst intensity and duration	0.079	7.311	2.80	0.001**
		1	Burst intensity and duration	0.103	4.661	2.81	0.012*
		2	Onset type	0.464	85.804	1.80	0***
	Suppression	1	Onset type	0.492	79.412	1.82	0***
		2	Burst intensity and duration	0.027	2.207	2.80	0.117
		1	Burst intensity and duration	0.041	1.715	2.81	0.186
		2	Onset type	0.478	79.411	1.80	0***

In each experiment, the analysis is conducted separately for the suppression and control conditions
†-p<0.1; *-p<0.05; **-p<0.01; ***-p<0.001

2. The Effect of Lexical Analogy

Our conclusions so far suggest that the misidentification of ill-formed syllables is not solely due to auditory/phonetic difficulties. However, syllables like *lbif* could be prone to misidentification due to lexical (i.e., non-grammatical) sources. Specifically, people might misidentify *lbif* because it does not resemble any other syllables stored in their lexicon.

To evaluate this lexical explanation, we next asked whether the effect of onset structure is subsumed by the lexical properties of our materials. To rule out the contribution of auditory/phonetic factors, we limited these analyses to the findings with printed materials (Experiments 3–4). We thus assessed several statistical properties of the printed materials, including the number of orthographic neighbors (i.e., the number of words obtained by substituting a single letter), the neighbors' summed frequency, the word's bigram count (i.e., number of words sharing two adjacent letters in the whole word) and its bigram frequency.[8] The summary statistics of these properties are provided in Table 7.

We contrasted the effect of these statistical measures and onset structure in a series of step-wise regression analyses. Given that the superior discrimination of sonority falls (e.g., *lbif*) clearly violates the *grammatical* account (most likely, due to a visual verification strategy discussed earlier), we limited the analyses to obstruent-initial syllables (e.g., *blif, bnif, bdif*). To measure the unique effect of statistical properties, we first forced onset type into the model and then entered the statistical properties (number of neighbors, neighbors' frequency,

[8] The neighborhood measures were obtained from the Speech and Hearing Lab Neighborhood Database (Nusbaum et al. 1984), and the bigram measures were based on Kučera and Francis (1967) database, excluding words that contain apostrophes, hyphens or spaces.

Table 7 The statistical properties of the materials used in Experiments 3–4

Onset type	Number of neighbors		Neighbors' frequency		Bigram count		Bigram frequency	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
blif	3	3.04	59	93.93	35	29.69	1288	1539.53
bnif	0	1.04	21	80.77	13	12.99	380	447.83
bdif	1	1.73	42	107.62	11	10.59	325	396.15

bigram count, and bigram frequency) together as the last predictor. We next gauged the unique effect of onset type by reversing the order of predictors, with onset type entered in the last step. Because the effect of onset type obtained irrespective of suppression, these regression analyses were conducted across suppression conditions in both experiments. To further ensure that suppression did not affect the unique effect of onset type, we then repeated these analyses under the suppression condition only. All analyses used sensitivity (d') as the dependent measure.

Results from both experiments (Table 8) showed that statistical properties did not uniquely capture participants' behavior. By contrast, the unique effect of onset type was significant even after controlling for the contribution of statistical properties. Critically, the unique effect of onset type was significant in both experiments even under the suppression condition.

These results converge with previous findings, suggesting that the syllable hierarchy cannot be explained only by the statistical properties of the English lexicon (e.g., Berent 2008; Berent and Lennertz 2010; Lennertz 2010). The fact that these conclusions obtain with printed materials, and even under suppression, further challenges both the auditory/phonetic and the

Table 8 The unique effect of (A) statistical properties (number of orthographic neighbors, neighbors' frequency, bigram count and frequency of the whole word); and (B) onset type in step-wise regression analyses in Experiments 3–4

Experiment	Condition	Step	Predictor	ΔR^2	ΔF	*df*	*P* value
3	Across suppression conditions	1	Onset type	0.123	11.464	1.82	0.001**
		2	Statistical properties	0.077	1.88	4.78	0.122
		1	Statistical properties	0.158	3.715	4.79	0.008**
		2	Onset type	0.041	4.044	1.78	0.048*
	Suppression	1	Onset type	0.17	16.811	1.82	0***
		2	Statistical properties	0.082	2.125	4.78	0.086†
		1	Statistical properties	0.199	4.916	4.79	0.001**
		2	Onset type	0.052	5.46	1.78	0.022*
4	Across suppression conditions	1	Onset type	0.165	16.169	1.82	0***
		2	Statistical properties	0.037	0.902	4.78	0.467
		1	Statistical properties	0.128	2.896	4.79	0.027*
		2	Onset type	0.074	7.204	1.78	0.009**
	Suppression	1	Onset type	0.074	6.588	1.82	0.012*
		2	Statistical properties	0.048	1.067	4.78	0.379
		1	Statistical properties	0.062	1.315	4.79	0.272
		2	Onset type	0.06	5.33	1.78	0.024*

Data is comprised of d' responses to the obstruent-initial syllables (e.g., blif, bnif, bdif) *motor embodiment* accounts. Together, these findings suggest that the restriction on onset structure might emanate from an abstract grammatical source.

Conclusion

This research examined whether the restrictions on syllable structure reflect abstract linguistic knowledge, or whether they are embodied in sensory and motor constraints. Contrary to the sensory embodiment account, we found that the effect of syllable structure obtains irrespective of stimulus modality (auditory or printed materials), and it is inexplicable by either phonetic cues or the orthographic similarity of the materials to English words.

Our results do not support the articulatory embodiment account either. We found that articulatory suppression clearly impaired participants' overall performance, a finding that could reflect the contribution of articulatory simulation to perception. Critically, suppression spared the effect of onset type. These results are consistent with the previous TMS findings of Berent and colleagues (Berent et al. 2015). As in our present study, the disruption to the articulatory motor system by TMS did impair overall sensitivity (d') in the syllable count task, but participants' sensitivity to the syllable hierarchy remained intact even under TMS.

The convergence between our behavioral experiments and the TMS research is of great importance because methodologically, they complement each other. Indeed, each method of suppression exhibits both advantages and disadvantages. TMS has the advantage of targeting specific motor sites without imposing additional attentional demands that are likely associated with mechanical suppression. However, mechanical suppression can address several potential limitations of the previous findings from TMS (Berent et al. 2015). First, TMS disruption might not suppress the articulator of interest fully, and it typically targets only a single articulator at a time (e.g., either tongue or lip but not both). Moreover, TMS effects may not be selective, as the electromagnetic pulses could also disrupt adjacent cortical regions that are irrelevant to articulatory motor control. Most critically, since the link between brain anatomy and cognitive function is rather tentative, the functional interpretation of TMS results is not entirely clear. Specifically, the finding that motor areas, for instance, play a role in speech production does not rule out the possibility that these areas might also mediate non-articulatory functions, including functions that are critical to grammatical computations. If so, the stimulation of motor regions could extend to non-motor, or even grammatical functions.

The convergence between our present results and the previous TMS findings is thus significant. The results from both studies suggest that English speakers possess broad grammatical preferences that are irreducible to articulatory simulation, statistical properties or auditory cues. Articulatory simulation might well contribute to speech perception, but it does not subsume the grammatical effect of syllable structure.

Compliance with Ethical Standards

Conflict of interest The authors declare no conflict of interest.

Ethical Standards This research was approved by the Institutional Review Board at Northeastern University.

Appendices

Appendix A: Monosyllabic Nonwords in Experimental Lists 1 and 2

List 1			
Large sonority rise	Small sonority rise	Sonority plateau	Sonority fall
blif	bwif	bdif	lbif
clim	cnim	cpim	lpim
drif	dlif	dbif	rdif
dwib	dmip	dgip	mdip
drip	dnup	dbup	rdup
glon	gmon	gbon	ifon
gref	gmef	gbef	rgef
kraf	kmaf	kpaf	rgaf
clop	cmup	ctop	ltop
cref	cmep	ctep	rkep
praf	pnaf	ptaf	rpaf
trok	tnok	tkok	rtok
twaf	tmaf	tpaf	mtaf
twuk	tnuk	tguk	mguk
twog	tmok	tpok	mtok

List 2			
Large sonority rise	Small sonority rise	Sonority plateau	Sonority fall
brop	bnop	bdop	rgop
crek	cnek	cteg	rtek
drof	dlof	dgof	rdof
dwup	dmup	dgup	mdup
drish	dnish	dgish	rbish
glep	gmep	gdep	lgep
gwid	gmit	gbit	mgit
klef	kmef	ktef	lkef
crik	cnik	ctig	rkik
cwug	cnuk	cpok	mcuk
plik	pnik	pkik	ltik
truf	tluf	tkuf	rtuf
twep	tlep	tkep	mtep
tref	tnef	tpif	rtef
trap	tmap	tpap	rpap

Appendix B: Monosyllabic Nonwords used in Experiments 1–4

Large sonority rise	Small sonority rise	Sonority plateau	Sonority fall
blif	bwif	bdif	lbif
brɑp	bnɑp	bdɑp	rgɑp
klim	knim	kpim	lpim
krɛk	knɛk	ktɛg	rtɛk
drif	dlif	dbif	rdif
drɑf	dlɑf	dgɑf	rdɑf
dwip	dmip	dgip	mdip
dwʊp	dmʊp	dgʊp	mdʊp
drʊp	dnʊp	dbʊp	rdʊp
driʃ	dniʃ	dgiʃ	rbiʃ
glɛp	gmɛp	gdɛp	lgɛp
glɑn	gmɑn	gbɑn	lfɑn
grɛf	gmɛf	gbɛf	rgɛf
gwit	gmit	gbit	mgit
klɛf	kmɛf	ktɛf	lkɛf
kræf	kmæf	kpæf	rgæf
krik	knik	ktig	rkik
kwʊg	knʊk	kpɑk	mkʊk
klɑp	kmʊp	ktɑp	ltɑp
krɛp	kmɛp	ktɛp	rkɛp
plik	pnik	pkik	ltik
præf	pnæf	ptæf	rpæf
trʊf	tlʊf	tkʊf	rtʊf
twɛp	tlɛp	tkɛp	mtɛp
trɑk	tnɑk	tkɑk	rtɑk
twæf	tmæf	tpæf	mtæf
trɛf	tnɛf	tpif	rtɛf
twʊk	tnʊk	tgʊk	mgʊk
træp	tmæp	tpæp	rpæp
twɑg	tmɑk	tpɑk	mtɑk

References

Abler, W. L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Systems*, *12*(1), 1–13.

Acuna, B. D., Eliassen, J. C., Donoghue, J. P., & Sanes, J. N. (2002). Frontal and parietal lobe activation during transitive inference in humans. *Cerebral Cortex*, *12*(12), 1312–1321.

Baddeley, A., Thompson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575–589.

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 and Performance*, *34*(5), 1288–1304.

Berent, I. (2013). The phonological mind. *Trends in Cognitive Sciences*, *17*(7), 319–327.

Berent, I., Brem, A., Zhao, X., Seligson, E., Pan, H., Epstein, J., et al. (2015). Role of the motor system in language knowledge. *Proceedings of the National Academy of Sciences*, *112*(7), 1983–1988. doi:10.1073/pnas.1416851112.

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., Lennertz, T., & Balaban, E. (2012a). Language universals and misidentification: A two way street. *Language and Speech*, *55*(3), 311–330.

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., & Rosselli, M. (2012b). Universal linguistic pressures and their solutions: Evidence from Spanish. *The Mental Lexicon*, *7*(3), 275–305.

Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, *26*(1), 75–108.

Berent, I., Pan, H., Zhao, X., Epstein, J., Bennett, M. L., Deshpande, V., et al. (2014). Language universals engage Broca's area. *PLoS ONE*, *9*(4), e95155. doi:10.1371/journal.pone.0095155.

Berent, I., & Perfetti, C. (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. *Cognition*, *104*(3), 591–630.

Berent, I., Vaknin-Nusbaum, V., Balaban, E., & Galaburda, A. M. (2013). Phonological generalizations in dyslexia: The phonological grammar may not be impaired. *Cognitive Neuropsychology*, *30*(5), 285–310.

Berent, I., Zhao, X., Balaban, E., & Galaburda, A. (2016). Phonology and phonetics dissociate in dyslexia: Evidence from adult English speakers. *Language, Cognition and Neuroscience*, *31*(9), 1178–1192.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *The Quarterly Journal of Experimental Psychology*, *39*(3), 467–478.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *36*(4), 701–711.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. New York, NY: Cambridge University Press.

Bolognini, N., & Ro, T. (2010). Transcranial magnetic stimulation: Disrupting neural activity to alter and assess brain function. *The Journal of Neuroscience*, *30*(29), 9647–9650.

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1508631112.

Bybee, J. (2008). Formal universals as emergent phenomena: The origins of structure preservation. In *Linguistic universals and language change* (pp. 108–121). Oxford University Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (2005). The role of features in phonological inventories. In C. C. E. Raimy (Ed.), *Contemporary views on architecture and representations in phonological theory* (pp. 19–68). Cambridge, MA: MIT Press.

D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, *48*(7), 882–887.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, *19*(5), 381–385.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, *55*, 149–179.

Eiter, B. M., & Inhoff, A. W. (2008). Visual word recognition is accompanied by covert articulation: Evidence for a speech-like phonological representation. *Psychological Research*, *72*(6), 666–674.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–448.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*(2), 399–402.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, *14*(1), 3–28.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361–377.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, *42*(3), 237–288.

Glenberg, A. M., Havas, D., Becker, R., & Rinck, M. (2005). Grounding language in bodily states: the case for emotion. In R. Zwaan & D. Pecher (Eds.), *The grounding of cognition: The role of perception and action in memory, language, and thinking* (pp. 115–128). Cambridge: Cambridge University Press.

Gómez, R. M., Berent, I., Benavides-Varela, S., Bion, R. A. H., Cattarossi, L., Nespor, M., et al. (2014). Language universal at birth. *Proceedings of the National Academy of Sciences*, *111*(6), 5387–5841.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant sequences. In E. A. Moravcsik (Ed.), *Universals of human language* (Vol. 2, pp. 243–279). Stanford, CA: Stanford University Press.

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics*, *17*(4), 315–339.

Hatsopoulos, N. G., & Suminski, A. J. (2011). Sensing with the motor cortex. *Neuron*, *72*(3), 477–487.
Hitch, G., & Baddeley, A. (1976). Verbal reasoning and working memory. *The Quarterly Journal of Experimental Psychology*, *28*(4), 603–621.
Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, *20*(2), 219–274.
Kleiman, G. M. (1975). Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior*, *14*(4), 323–339.
Kučera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Dartmouth Publishing Group.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic books.
Lee, A. B., Mumford, D., & Huang, J. (2001). Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision*, *41*(1/2), 35–59.
Lennertz, T. (2010). *People's knowledge of phonological universals: Evidence from fricatives and stops*. Unpublished doctoral dissertation, Northeastern University.
Lennertz, T., & Berent, I. (2015). On the sonority levels of fricatives and stops. *The Mental Lexicon*, *10*(1), 88–132.
Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*(2), 123–132.
Levy, B. A. (1975). Vocalization and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(3), 304–316.
Levy, B. A. (1978). Speech processing during reading. In A. Lesgold, J. Pellegrino, S. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction*. New York, NY: Plenum Publishing Corporation.
Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36.
Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research*, *66*(4), 579–619.
Lukatela, G., Eaton, T., Sabadini, L., & Turvey, M. (2004). Vowel duration affects visual word identification: Evidence that the mediating phonology is phonetically informed. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 151–162.
MacNeilage, P. (2008). *The origin of speech*. New York, NY: Oxford University Press.
Maïonchi-Pino, N., de Cara, B., Ecalle, J., & Magnan, A. (2012). Are french dyslexic children sensitive to consonant sonority in segmentation strategies? Preliminary evidence from a letter detection task. *Research in Developmental Disabilities*, *33*(1), 12–23.
Martin, M. (1978). Speech recoding in silent reading. *Memory & Cognition*, *6*(2), 108–114.
Martin, C. G., & Games, P. A. (1977). ANOVA tests for homogeneity of variance: Nonnormality and unequal samples. *Journal of Educational statistics*, *2*(3), 187–206.
Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *The Journal of Neuroscience*, *29*(31), 9819–9825.
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*(10), 357–376.
O'Shea, J., & Walsh, V. (2007). Transcranial magnetic stimulation. *Current Biology*, *17*(6), R196–R199.
Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, *36*(1), 55–90.
Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar.* Wiley-Blackwell.
Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, *103*(20), 7865–7870. doi:10.1073/pnas.0509989103.
Samuel, A. G. (2011). Speech perception. *Annual Review of Psychology*, *62*(16), 1–24.
Sanes, J. N., & Donoghue, J. P. (2000). Plasticity and primary motor cortex. *Annual Review of Neuroscience*, *23*(1), 393–415.
Schwartz, J.-L., Abry, C., Boë, L.-J., & Cathiard, M. (2002). Phonology in a theory of perception-for-action-control. In J. Durand & B. Laks (Eds.), *Phonetics, phonology, and cognition* (pp. 254–280). New York, NY: Oxford University Press.
Tamási, K., & Berent, I. (2014). Sensitivity to Phonological Universals: The case of stops and fricatives. *Journal of Psycholinguistic Research*, *44*(4), 359–381.
Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*(4), 488–522.

Wilson, C., & Davidson, L. (2013). Bayesian analysis of non-native cluster production. In *Proceedings of NELS* (Vol. 40).

Wilson, C., Davidson, L., & Martin, S. (2014). Effects of acoustic-phonetic detail on cross-language speech production. *Journal of Memory and Language*, *77*, 1–24.

Zhao, X., & Berent, I. (2015). Universal restrictions on syllable structure: Evidence from Mandarin Chinese. *Journal of Psycholinguistic Research*. doi:10.1007/s10936-015-9375-1.