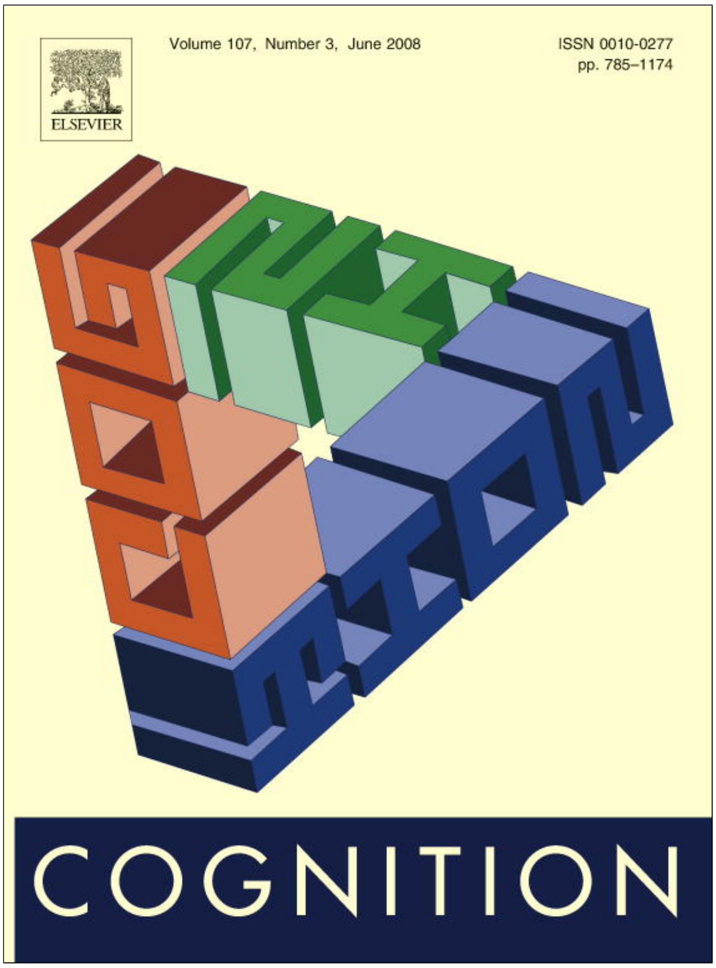
Volume 107, Number 3, June 2008
ISSN 0010-0277
pp. 785–1174
ELSEVIER
COGNITION

Available online at www.sciencedirect.com
ScienceDirect

Cognition 107 (2008) 785–816

COGNITION

www.elsevier.com/locate/COGNIT

Production constraints on learning novel onset phonotactics

Melissa A. Redford *

The University of Oregon, Linguistics Department, 1290 University of Oregon, Eugene, OR 97403, USA

Received 2 January 2007; revised 15 November 2007; accepted 24 November 2007

Abstract

Three experiments addressed the hypothesis that production factors constrain phonotactic learning in adult English speakers, and that this constraint gives rise to a markedness effect on learning. In Experiment 1, an acoustic measure was used to assess consonant–consonant coarticulation in naturally produced nonwords, which were then used as stimuli in a phonotactic learning experiment. Results indicated that sonority-rising sequences were more coarticulated than -plateauing sequences, and that listeners learned novel-rising onsets more readily than novel-plateauing onsets. Experiments 2 and 3 addressed the specific questions of whether (1) the acoustic correlates of coarticulation or (2) the coarticulatory patterns of self-productions constrained learning. In Experiment 2, stimuli acoustics were altered to control for coarticulatory differences between sequence type, but a clear markedness effect was still observed. In Experiment 3, listeners' self-productions were gathered and used to predict their treatment of novel-rising and -plateauing sequences. Results were that listeners' coarticulatory patterns predicted their treatment of novel sequences. Overall, the findings suggest that the powerful effects of statistical learning are moderated by the perception–production loop in language.

Keywords: Statistical learning; Phonotactics; Markedness; Syllabification; Phonetic constraints; Coarticulation

* Tel.: +1 541 346 3789; fax: +1 541 346 5961.
E-mail address: redford@uoregon.edu

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2007.11.014

1. Introduction

Language specific knowledge about phoneme co-occurrences informs word segmentation (McQueen, 1998; Norris, McQueen, Cutler, & Butterfield, 1997) and syllabification judgments (Redford & Randall, 2005; Smith & Pitt, 1999; Treiman & Zukowski, 1990). For example, English-speaking listeners are aware that English words and syllables cannot begin with a *tl* sequence, even though the sequence can occur word-internally. In phonological terms, the *tl* sequence is an illegal onset. English speakers know this, and the result is that they syllabify words such as Atlántic so that *tl* forms an offset–onset sequence (At-lán-tic) rather than an onset cluster to the second syllable (cf. a-tró-cious).

The cumulative knowledge about legal and illegal onsets and offsets in a language is referred to as phonotactic knowledge. This is one aspect of phonological knowledge that is indisputably learned, given that each language has different phonotactics. Further, recent experiments with both infants and adults indicate that novel phonotactic patterns are abstracted very quickly over a set of data (Chambers, Onishi, & Fisher, 2003; Dell, Reed, Adams, & Meyer, 2000; Onishi, Chambers, & Fisher, 2002; Warker & Dell, 2006). For example, Dell et al. (2000) found that novel phonotactic constraints were reflected in experimentally induced speech errors during the first test session. Similarly, Onishi et al. (2002) found that adults learned novel syllable position and sequencing restrictions from just a few minutes worth of auditory exposure to stimuli that embodied these phonotactic restrictions.

The results on phonotactic learning and on statistical learning of language more generally suggest that linguistic knowledge is dynamic – subject to change with new perceptual input (Pierrehumbert, 2003; Fisher, Church, & Chambers, 2004). Such a suggestion is consistent with the facts of language change, and especially with changes in individual language behavior (e.g., Harrington, Palethorpe, & Watson, 2000). But the sheer power of statistical learning seems to over predict the rate of language change. If only a few minutes of exposure to novel language stimuli is sufficient for abstracting new language knowledge, why is individual language behavior as stable as it is in the face of variable input? For that matter, why is language as stable as it is given high inter-speaker variability? Although languages may continually change, certain patterns are stable over time and highly frequent across languages. How is this possible, if linguistic knowledge is is dynamic?

A common answer to these questions is that statistical learning is constrained (e.g., Newport & Aslin, 2000; Saffran, 2003). In particular, cognitive and perceptual-motor factors are assumed to constrain language learning. These same factors are also thought to shape the emergence and maintenance of unmarked language patterns – patterns that are stable and cross-linguistically frequent. Unstable and isolated, or marked, patterns are assumed to emerge through historical accident and to represent only transitional structures in a language's history.

If statistical language learning is indeed constrained by the same factors that give rise to unmarked language patterns, then these patterns should be easier to learn than marked patterns. The first aim of the present study was to test this hypothesis for phonotactic learning in English-speaking adults. The second aim was to account for any differential learning of unmarked and marked sound patterns in terms of more basic constraints, namely, in terms of the perceptual-motor factors from which these patterns are assumed to arise.

1.1. Markedness and phonotactic learning

Unmarked and marked phonotactic patterns are defined with reference to the Sonority Sequencing Principle (SSP), which states that segments are sequenced so that sonority increases from the edges of a syllable to its center (Clements, 1990; Hooper, 1972; Selkirk, 1982). The corollary to the SSP is that greater sonority differences between segments are preferred to smaller sonority differences. For onset clusters, this means that the SSP describes a hierarchy of preferred segment sequencing in which obstruent-liquid onsets, like *sl*-, are better than obstruent-nasal onsets, like *sn*- (assuming a basic sonority scale obstruents < nasals < liquids < glides < vowels). It also means that obstruent–obstruent sequences that violate the SSP, like *sp*-, are disfavored. Although the preference hierarchy also parallels the typological data (Bell & Hooper, 1978), the principal markedness contrast is between the cross-linguistically frequent sonority-rising sequences and the cross-linguistically infrequent sonority-plateauing or sonority-reversing sequences.

The first question addressed in the current study was whether cross-linguistically infrequent sonority-plateauing onset sequences are learned as readily as cross-linguistically frequent sonority-rising onset sequences by native English speakers. Adult listeners were exposed to naturally produced sonority-rising and sonority-plateauing word onset consonant sequences that do not occur in English, and then asked to syllabify the same sequences in word medial position. The expectation was that listeners would learn that the novel word onsets were possible (i.e., legal) onset clusters, and that this learning would generalize to influence syllabification behavior: listeners would more often syllabify novel word medial sequences as onset clusters to the second syllable when they had been exposed to these sequences in word onset position than when they had not. If markedness constrains phonotactic learning, then listeners would be more likely to generalize in this way from novel sonority-rising word onsets than from novel sonority-plateauing word onsets.

The specific focus on learning novel onset clusters makes the present study different from previous studies on phonotactic learning, which have typically focused on distributional patterns (e.g., *η* is restricted to syllable-offset position in English) rather than on the identity of a single structure (e.g., *tl*- is not a possible onset in English). However, the hypothesis that markedness constrains learning has been investigated in several recent studies on phonological learning in adults and infants. The results of these studies tend to support the hypothesis (e.g., Newport & Aslin, 2004; Peperkamp, Skoruppa, & Dupoux, 2006; Wilson, 2006; but see Seidl & Buckley, 2005). For example, Newport and Aslin (2004) found that listeners had trouble extracting pseudo-words from a speech stream given nonadjacent dependencies between syllables, consistent with the typological fact that such dependencies are cross-linguistically rare. They also found that listeners easily learned to extract pseudo-words from a speech stream given nonadjacent dependencies between segments, consistent with the typological fact that such dependencies are cross-linguistically common.

Newport and Aslin (2004) attributed the differential learning of segmental and syllabic dependencies to a statistical learning mechanism that computes one type of dependency more readily than the other, and offered several explanations for why this would be. All of the explanations grounded the difference in lower-level processing, namely, in speech perception. This type of constraint grounding also figures in studies that have investigated markedness constraints on phonological rule learning (e.g., Peperkamp et al., 2006; Wilson, 2006). For example, Wilson (2006) investigated the degree to which adult listeners would generalize a pattern of velar palatalization to a new vowel context, and found that generalized learning depended on the extent to which the novel vowel context predisposes the production and perception of palatalization. Listeners were more likely to generalize palatalization from a mid-front vowel context to a high-front vowel context than they were to generalize from a high-front vowel context to a mid-front vowel context. Since the particular direction of generalization matched the implicational relationships found across languages, Wilson argued that phonetic substance must be encoded in the phonological constraints that are acquired with grammars.

Overall, then, studies that document constraints on phonological learning typically ground the constraints in more basic phonetic processes. This practice suggests that the best explanation for markedness effects is a mechanistic one. Accordingly, the next section develops a phonetic basis for the SSP and alternative hypotheses for how the phonetics might influence the acquisition of new syllable onsets. The goal is to provide a more complete rationale for the prediction that listeners will be more likely to generalize syllable-onset structures from novel sonority-rising word onsets than from novel sonority-plateauing word onsets.

1.2. Phonotactic markedness and phonetic naturalness

The practice of explaining markedness effects in terms of lower-level processes is consistent with the now mainstream view that cross-language patterns, like those described by the SSP, are universal or highly preferred because they are phonetically motivated (e.g., Blevins, 2004; Demolin & Soquet, 1999; Hayes, Kircher, & Steriade, 2004; Hume & Johnson, 2001; Kingston & Diehl, 1994; Lindblom, MacNeilage, & Studdert-Kennedy, 1984; Ohala, 1993; Pierrehumbert, 2001; Redford, Chen, & Miikkulainen, 2001).[1] In this section, some previous work on the phonetics of preferred segment sequencing is synthesized to ground the SSP in perceptual-motor processes with an emphasis on speech production.

A prominent view in phonetics is that the preferred segment sequencing patterns described by the SSP emerge across languages from coarticulatory patterns in produc-

[1] The references to Hume and Johnson (2001) and to Hayes et al. (2004) are references to edited volumes. Both volumes contain contributions from linguists who advance phonetically based explanations for phonological patterns that occur across the world's languages (e.g., assimilation, metathesis, lenition). Almost all of the phonetic explanations advanced are perceptually based.

tion and their perceptual effects (Davidson, 2006a; Kawasaki-Fukumori, 1992; Lindblom, 1983; Ohala & Kawasaki-Fukumori, 1997; Steriade, 1999; Wright, 2004). A specific proposal that follows from work on the relationship between prosodic structure and jaw movement (de Jong, Beckman, & Edwards, 1993; Erickson, 1998; Erickson, Fujimura, & Pardon, 1998; Harrington, Fletcher, & Roberts, 1995; Stone, 1981) as well as from work on the relationship between sound sequencing and jaw movement (MacNeilage, 1998; MacNeilage & Davis, 2000; MacNeilage, Davis, Kinney, & Matyear, 2000) is that sequences that are easily coarticulated within the open–close cycle of the jaw are grouped together for speech output. Grouping segments in this way allows a speaker to work with the biomechanical dependencies between the lips, tongue, and jaw instead of against them. Segment sequences that are coarticulated within a single jaw cycle parallel the preferred within syllable sequencing patterns described by the SSP because the sonority hierarchy parallels an articulatory openness hierarchy (Browman & Goldstein, 1989; Butt, 1992; Lindblom, 1983). Onset sequences that violate the SSP are less amenable to grouping according to a single continuous open–close movement of the jaw and so are coarticulated at a perceptual cost. For example, consonantal release (opening) gestures, which provide information about segment identity, are minimized or lost in plateauing sequences that are fully coarticulated (Davidson, 2006a; Steriade, 1999; Wright, 2004).

Thus, the proposal is that the SSP follows from a default chunking strategy in production and from the perceptual costs that ensue from coarticulating segments that require the same degree of vocal tract constriction (i.e., jaw height). Significantly, coarticulatory groupings of more and less sonorous elements in speech have long been linked to syllabic representations (see Krakow, 1999, for a review), which encode language-specific phonotactics. If we assume a bottom-up directionality to the link between coarticulatory grouping and syllabic representation, a specific hypothesis regarding phonotactic learning follows: perceptual-motor processes drive initial coarticulatory groupings, thereby constraining syllabic representations, which in turn constrain the acquisition of novel phonotactics. With regard to learning novel onsets, this hypothesis predicts that language users will be more likely to identify prevocalic consonant sequences as homosyllabic if they are coarticulated and overlapped with the subsequent vowel during production than if they are not (cf. Browman & Goldstein, 1990). The user may infer that sequences are produced in this way either by referencing the acoustic correlates of coarticulation available in the signal or by simulating the production of a sequence to reference one's own coarticulatory patterns. In either case, users will be more likely to extract novel sonority-rising onset clusters than novel sonority-plateauing onset clusters since the former are assumed to be more easily coarticulated with a subsequent vowel than the latter.

1.3. The current study

Again, the first aim of the current study was to test the hypothesis that markedness constrains phonotactic learning. Listeners were tested on their ability to extract possible onset clusters from novel sonority-rising and sonority-plateauing word onset sequences respectively, using word medial syllabification as a measure of phonotactic learning.

The prediction was that English listeners would more readily learn novel onset phonotactics from unmarked sonority-rising word-onset sequences than from equally novel marked sonority-plateauing sequences. Such a prediction is motivated by the hypothesis that sonority-rising sequences are more easily coarticulated in production and so more readily identified as homosyllabic clusters than sonority-plateauing sequences.

The second aim of the current study was to test the more basic hypothesis that coarticulation constrains phonotactic learning. Two alternatives were investigated: (1) listeners extract an onset cluster from the signal when the acoustics indicate that a consonant sequence is produced with substantial articulatory overlap; (2) listeners reference their own productions of a consonant sequence to extract phonological structure, and extract an onset cluster when they coarticulate the sequence and overlap it with the following vowel.

The first and second aim of the study were addressed in parallel in Experiment 1. Naturally produced nonwords were used as stimuli in a phonotactic learning experiment, and an acoustic measure was used to assess the degree of consonant–consonant coarticulation for the different stimulus types. Results on the learning of novel-rising and novel-plateauing onsets were then considered in the context of the different acoustic patterns on consonant–consonant coarticulation. Experiments 2 and 3 addressed the more specific questions of whether it is (1) the acoustic correlates of coarticulation or (2) the coarticulatory patterns of self-productions that influence phonotactic learning. In Experiment 2, the stimuli acoustics were altered so that a correlate of coarticulation and cue to boundary location was controlled during the phonotactic learning experiment. The goal was to test whether all novel onsets would be equally well learned when coarticulatory differences were absent in the input. In Experiment 3, listeners' self-productions were gathered to determine whether these predicted learning of novel-rising and novel-plateauing onset clusters better than the stimulus acoustics.

2. Experiment 1: Phonotactic learning of marked and unmarked sequences

If listeners readily learn novel phonotactics, then they should be willing to expand their notion of possible onset cluster with minimal exposure to novel word-edge consonant sequences. This expectation follows from both linguistic theory and psycholinguistic work, which indicates that word-edge patterns influence syllable structure representations (Berg & Niemi, 2000; Hooper, 1972; Pulgram, 1970; Redford & Randall, 2005; Selkirk, 1982; Smith & Pitt, 1999; Treiman & Zukowski, 1990). Outside the laboratory, listeners are likely to engage in such learning when exposed to a new language that has different phonotactics from the native language. In such a situation, listeners would have access to all the information that is available in natural speech – information that is presumably relevant to extracting linguistic structure. To replicate this likely real world scenario within the parameters of an experiment, English listeners were exposed to stimuli produced by native Russian speakers.

Russian has extremely permissive word initial phonotactics. Many Russian word onsets are also possible English onsets, but many more are not. I used this fact about Russian to construct control and test sets of stimuli with familiar and novel sequences in word onset and word medial position. The novel sequences either conformed to or violated the SSP, that is, they were either sonority-rising or sonority-plateauing sequences. Several native Russian speakers produced the nonsense words, which were then presented to English-speaking listeners during a phonotactic learning experiment. But first, stimuli acoustics were measured to determine the degree to which Russian speakers coarticulated the consonants in the familiar, novel-rising, and novel-plateauing sequence types. In particular, the duration ratio of the two consonants was calculated to quantify the consonantal duration pattern that arises from consonant–consonant coarticulation and overlap with a subsequent vowel.

Although the consonantal duration pattern does not perfectly reflect the underlying kinematics of coarticulation because kinematics-to-acoustics is a many-to-many mapping (cf. Cho, 2004), the pattern is (a) present across all sequence types, (b) available to listeners, and (c) provides a cue to syllable structure (Christie, 1977; Redford & Randall, 2005; Tuller & Kelso, 1991). In particular, the cumulative research indicates that onset clusters and offset–onset sequences are produced with distinct patterns. Syllable-onset clusters are produced with a long-short durational pattern (e.g., Christie, 1977; Haggard, 1973), which emerges when the acoustical duration of the second consonant is truncated by carry-over effects from the first consonant and by anticipatory effects of the following vowel (Redford, 2007). An offset–onset pattern is produced with either a long–long or short–long pattern (e.g., Anderson & Port, 1994; Christie, 1977), which emerges when consonant–consonant coarticulation is blocked by a word or phrase boundary or by sequential articulatory factors (Fougeron & Keating, 1997; Redford, 2007).

The stimuli in the present study controlled for the prosodic factors of lexical stress and word position, but speakers were nonetheless expected to produce the rising and plateauing sequences with different durational patterns. This expectation followed from the assumption that perceptual-motor factors allow speakers to group rising sequences with the following vowel more readily than plateauing sequences. Thus, the expectation was that Russian speakers would produce rising-type stimuli with a long–short durational pattern and plateauing-type stimuli with a long–long or short–long pattern.

Once the acoustics were measured, native English-speaking listeners were exposed to the stimuli with either familiar or novel word onset sequences during training and asked to syllabify both familiar and novel word medial sequences during test. A training-dependent increase in the number of onset cluster syllabifications of novel medial sequences would indicate that listeners had generalized the novel phonotactics from word onset sequences. Listeners were expected to generalize from and to onset sequences with long–short durational patterns more readily than they would generalize from or to those sequences with long–long or short–long patterns.

2.1. Method

2.1.1. Participants

Forty-four undergraduate students from the University of Oregon participated in the phonotactic learning experiment in exchange for course credit. All were native American-English speakers and all reported normal hearing.

2.1.2. Stimuli

Four native Russian speakers – two male and two female – produced the stimuli. The speakers were financially compensated for their time. All stimuli were possible disyllabic words in Russian with either word onset consonant sequences (CCVCVC), word medial sequences (CVCCVC), or no consonant sequences (CVCVC). The word onset sequences were either consistent with English phonotactics (e.g., *drevat*) or not. The sequences that were not consistent with English phonotactics were either sonority-rising (e.g., *tlevat*) or sonority-plateauing (e.g., *bdevat*) sequences. All sequences also occurred in word medial position in the CVCCVC stimuli. Specifically, the pre-vocalic consonants and accompanying vowels in the CCVCVC words were inverted to create the CVCCVC words. For example, the nonsense word *drevat* yielded the corresponding nonsense word *vadret*. The stimuli with alternating consonants and vowels were created by deleting either the first or second consonant of the sequence (e.g., *drevat* → *devat*, *revat*; *vadret* → *vadet*, *varet*) and also by transposing the vowels (e.g., *devat* → *davet*; *revat* → *varet*).

The stimuli with consonant sequences were the target stimuli in the experiment. I will refer to those with sequences obeying English word-edge phonotactics as *familiar* (e.g., *drevat* and *vadret*) and those with sequences violating English word-edge phonotactics as *novel*. The novel stimuli with sonority-rising sequences (e.g., *tlevat* and *vatlet*) will be referred to as *novel-rising*, and those with sonority-plateauing sequences (e.g., *bdevat* and *vabdet*) as *novel-plateauing*. The target stimuli with word onset sequences were used in training and those with word medial sequences were used in testing, as described further below. The stimuli with alternating consonants and vowels (CVCVC) were filler stimuli, used to help obscure the relationship between training and testing.

There were eight different types of familiar, novel-rising, and novel-plateauing nonsense word stimuli (see Appendix). The differences were restricted to the consonant sequences, which were *bl*, *br*, *dr*, *fr*, *gl*, *gr*, *sl*, and *šr* in the familiar stimuli, *bn*, *dl*, *dn*, *gn*, *tl*, *vl*, *zr*, and *žn* in the novel-rising stimuli, and *bd*, *dv*, *mn*, *mr*, *tk*, *vs*, *zv*, and *žv* in the novel-plateauing stimuli. The final consonant was always *t* and the vowels were always either *a* or *e*. Altogether, then, there were 24 different CCVCVC nonsense words and 24 different CVCCVC words. These yielded 128 filler CVCVC words, which were reduced to 77 words once duplicate tokens were eliminated.

The stimuli were embedded in the frame sentence *Skaži* ___ *s nova* (Say ___ again), written in Cyrillic, and arranged as randomized lists, which were then given to the four Russian speakers. Each speaker was recorded in a sound attenuated booth as they read the list of phrases. The goal was to generate 48×4 target stimuli in addition to 77×4 filler stimuli. This goal was largely met, though it was later discovered that one speaker had missed one of the novel test words (*vadnet*).

Once recorded, the phrases were amplitude normalized across speakers to the level of *s nova* in the frame sentence. The consonant duration measures were taken at the same time. The criteria used to take these measures are described further below.

2.1.3. Acoustic measures

The stimuli were displayed in Praat (Boersma & Weenink, 2002) as oscillograms and spectrograms, and the acoustic durations of each of the consonants in the onset or medial sequence was measured, using standard segmentation procedures. The specific criteria for fricatives, stops, nasals, and liquids were as follows.

Fricative boundaries were defined by the sudden onset/offset of aperiodic energy. When two fricatives appeared in sequence, the boundaries between the two were distinguished by frequency and amplitude differences in the fricative waveform. For example, *v* was found to have a more periodic and lower amplitude waveform than either *s*, *z*, or *ž* with which it co-occurred. Stop boundaries were defined by a sudden drop/rise in the amplitude of the periodic waveform. Closure and release duration were summed to provide total stop duration. When stops co-occurred, the offset of the first stop's release was used as the onset boundary for the second stop. All speakers released all stops occurring in stop–stop sequences. The boundaries between nasals and vowels and between liquids and vowels were indicated by an abrupt change in waveform frequencies and amplitude. In the case of the lateral liquid, an abrupt boundary did not always occur. For these stimuli, the boundary was placed at the temporal midpoint of the transitioning 3rd formant (Klatt, 1976). When two nasals co-occurred, the boundary between the two was always marked by a brief release. Ambiguity in defining segment boundaries was resolved by repeated listening to different sections of the waveform, that is, by establishing a correspondence between spectral and amplitude changes and an auditory segmentation of the waveform.

The two consonant durations for a particular sequence were transformed into a C1 to C2 duration ratio (C1:C2 duration). The single value provided information about the consonantal duration pattern, the selected acoustic correlate of coarticulation and well-studied cue to boundary location. The value also provided a normalized measure, which was needed in order to compare the different types of sequences in the different word positions across the different speakers. The duration ratios were analyzed within speaker as a function of consonant sequence type (familiar, novel-rising, novel-plateauing) and of word position (onset vs. medial) in a repeated measures ANOVA. Paired *t*-tests were used to conduct post hoc mean comparisons, and alpha was adjusted downward with the number of comparisons.

The analysis indicated that the acoustic correlate of coarticulation varied significantly with sequence type [$F(2,30) = 7.66$, $p = .002$] and word position [$F(1,31) = 18.20$, $p < .001$]. The interaction between the two factors was also significant [$F(2,30) = 8.43$, $p = .001$]: the consonantal duration pattern varied more with word position in the familiar sequences than in the novel sequences. Fig. 1 shows this interaction.

In Fig. 1, the consonantal duration pattern is expressed on the *y*-axis as the ratio of consonant 1 duration to consonant 2 duration (C1:C2). If C1:C2 equaled 2, then the first consonant in the sequence was twice as long as the second. If C1:C2 equaled 1, then the consonants in the sequence were of equal duration. The figure shows that all sequences had C1:C2 duration ratios greater than 1 on average, indicating that C1 was typically longer than C2. The figure also shows that C1:C2 was greater in word

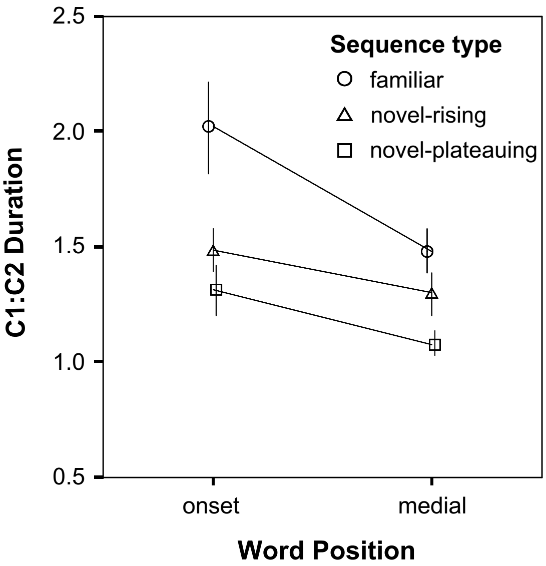

Fig. 1. The duration ratio of the first to second consonant is shown as a function of word position and sequence type for the Russian nonce word target stimuli.

onset position than in word medial position, consistent with a boundary effect that pushes consonants to be more tightly coarticulated and overlapped with the following vowel in word onset position than in word medial position.

The clear differences in C1:C2 between sequence type also indicate differences in the extent to which different consonant sequences were coarticulated. In fact, the pattern of results suggests a hierarchy of cohesiveness that parallels a goodness ranking according to sonority sequencing. C1:C2 duration was greatest for the obstruent-liquid sequences, which were the familiar onsets. C1:C2 duration was lower for the obstruent–sonorant (obstruent-liquid and obstruent-nasal) sequences that represented novel-rising onsets [familiar vs. rising; $t(63) = 2.93$, $p < .017$, $d = .56$]. Finally, C1:C2 duration was lowest for the obstruent–obstruent or sonorant–sonorant sequences that represented the novel-plateauing onsets [rising vs. plateauing; $t(63) = 2.47$, $p < .017$, $d = .42$]. Thus, the C1:C2 differences by sequence type paralleled the sonority differences, which were greatest for the familiar forms, lower for the novel-rising forms (liquids are more sonorous than nasals), and lowest for the novel-plateauing forms (no sonority difference). Assuming a link between coarticulation and syllable structure, the differences are consistent with the prediction that novel onset structure will be extracted more readily from novel-rising sequences than from novel-plateauing sequences. This prediction was tested in a phonotactic learning experiment, the procedures for which are described next.

2.1.4. Procedure

The phonotactic learning experiment included a training and testing phase. During the training phase, half of the listeners ($N = 23$) were assigned to the two control conditions in which they were presented with CCVCVC nonsense words that had familiar phonotactics (e.g., *drevat*). The other half ($N = 21$) were assigned to one of the two experimental conditions in which they were presented with CCVCVC words that had either novel-rising phonotactics (e.g., *tlevat*) or novel-plateauing phonotactics (e.g., *bdevat*). The CCVCVC stimuli were repeated twice ($N = 64$) and mixed in with half of the filler stimuli ($N = 154$) for a total of 218 stimuli. The CCVCVC target and CVCVC filler stimuli were then presented in pre-determined randomized orders over headphones to a maximum of two listeners, seated back-to-back at separate desks, in an experimental room. Listeners were able to control both the level of stimulus presentation and the rate at which they were presented. The level was adjusted using a volume control dial, and the rate by clicking on a button to advance to the next stimulus.

Listeners were told that the aim of the first part of the experiment was to familiarize them with Russian pronunciation and to give them a sense of how such words might be spelled. They were given a list of the different nonsense words they would be hearing written in normal English orthography, and were told to concentrate on how the written word was pronounced. The true goal was to minimize misperception during training by ensuring that listeners associated novel acoustic patterns with the sequences being produced.

Participants were told that each word would occur in a frame sentence, and that they were to ignore this; that the frame was there to make the speaking and listening task more natural. To help listeners concentrate on the spoken word and to ensure that they matched it with the written word, listeners were asked to syllabify the written nonsense words based on what they heard by placing a slash through the word on the paper. They were given the example of the English word *super* embedded in the Russian frame sentence *Skaži* ____ *s nova* and asked how they might break *super* apart. Once it was clear that the listeners understood the task, they were left alone to complete it.

Once listeners had completed the training phase, the experimenter set up the testing phase of the experiment. During this phase, both the control and experimental groups were presented with the familiar and novel sequences in word medial position (e.g., *vadret* and *vatlet* or *vadret* and *vabdet*). The four different train-test conditions are summarized in Table 1.

The target CVCCVC stimuli ($N = 64$) for the testing phase of the experiment were mixed with the other half of the filler CVCVC stimuli ($N = 154$), yielding the same ratio of target to filler items heard during training. Listeners were told that the object of this part of the experiment was to break apart the nonsense words they heard. They were given a blank sheet of paper and instructed to write down the word they heard and to indicate with a dash or a dot where they thought the word should be broken. This type of metalinguistic syllabification task has been used many times in the literature (e.g., Redford & Randall, 2005; Rubach & Booij, 1990; Treiman & Danis, 1988; Treiman & Zukowski, 1990). More importantly, such a task is known to elicit linguistic intuitions on syllable boundary location consistent with those that can be elicited by other less direct methods, such as infixing and word games (Treiman, 1983; Treiman & Danis, 1988).

Table 1
The control and experimental conditions are shown with the corresponding number of participants per condition

Test	Train	
	Familiar	Novel
Novel-rising	*drevat/vatlet* ($N = 11$)	*tlevat/vatlet* ($N = 12$)
Novel-plateauing	*drevat/vabdet* ($N = 11$)	*bdevat/vabdet* ($N = 10$)

2.1.5. Response coding

Syllabification judgments on the medial consonant sequences were categorized into three basic types: onset–offset judgments (VC.CV), onset cluster judgments (V.CCV), and the less common 'other' judgments, which included offset cluster judgments (VCC.V), singleton judgments (V.CV or VC.V), judgments with epenthetic vowels (V.CV.CV), and the occasional ambisyllabic judgment (VC.CCV). Listener judgments were characterized in terms of each of these three categories. For example, if a listener syllabified *vadret* as *va-dret* the judgment would be coded as follows: VC.CV = 0, V.CCV = 1, Other = 0. The coding scheme therefore captured the fact that syllabification judgments are categorical: one judgment precludes another.

2.2. Results and discussion

Analyses of boundary judgments indicated a significant training dependent increase in onset cluster syllabifications for listeners exposed to novel-rising sequences, but not for those exposed to novel-plateauing sequences. However, significant interactions between training and test type suggested that even listeners exposed to novel-plateauing sequences had learned some novel onset phonotactics. A regression analysis that used the consonantal duration ratio as a predictor variable indicated that only the most tightly coarticulated novel-plateauing sequences were treated as onset clusters. These overall results are presented in more detail below.

2.2.1. Effects of sequence type on phonotactic learning

The analyses on categorical syllabification judgments addressed two specific questions: (1) whether exposure to novel word onset sequences led to more onset cluster syllabifications of the same sequences in word medial position; and (2) whether the effect of training (and therefore learning) varied depending on sequence type. The analyses excluded those responses with too few or too many segments, that is, singleton judgments, ambisyllabic judgments, and judgments with epenthetic vowels. The remaining V.CCV, VC.CV, and V.CCV judgments were analyzed in a random effects logistic regression model using the GLIMMIX procedure. The within-subjects variable, test type (familiar or novel), was embedded within listener, a random effect. Listener was in turn embedded within training (familiar or novel) and within novel sequence type (rising or plateauing). Speaker was a between subjects factor that did not have a significant effect on listener judgments.

The overall fit of the model was good. The generalized χ^2 was 2169.60, yielding a generalized χ^2 to degrees of freedom ratio of 0.94 (</= 1 is a good fit; >1 is a poor fit, i.e., indicates overdispersion). A look at the predictor variables within each novel sequence type showed that test type strongly affected syllabification judgments in both the novel-rising and novel-plateauing conditions [rising, $B = 2.64$, $t(2278) = 15.11$, $p < .001$; plateauing, $B = 3.15$, $t(2278) = 15.28$, $p < .001$], but that training only significantly affected syllabification judgments in the novel-rising condition [rising, $B = -1.53$, $t(2278) = -2.63$, $p = .009$; plateauing, $B = -0.05$, $t(2278) = -0.08$, $p = .937$]. However, the interaction between training and test type was significant in both conditions [rising, $B = .35$, $t(2278) = 2.03$, $p < .042$; plateauing, $B = .44$, $t(2278) = 2.11$, $p = .034$].

Fig. 2 shows the direction of all within sequence type effects. For instance, the figure shows that listeners made more onset cluster judgments on sequences that obeyed word-edge phonotactics in English (familiar) than on those that did not (novel) regardless of training. This finding is consistent both with the well known effect of legality on syllabification (Smith & Pitt, 1999; Treiman & Zukowski, 1990) and with the acoustic result that Russian speakers produced these sequences with more overlap than the novel-rising or novel-plateauing sequences (see Fig. 1).

Fig. 2 also shows that listeners exposed to novel-rising word onset sequences during training made more onset cluster judgments overall than those exposed only to familiar onset sequences. This effect of training was not significant in the novel-plateauing condition, suggesting that listeners exposed to the novel-rising word onset

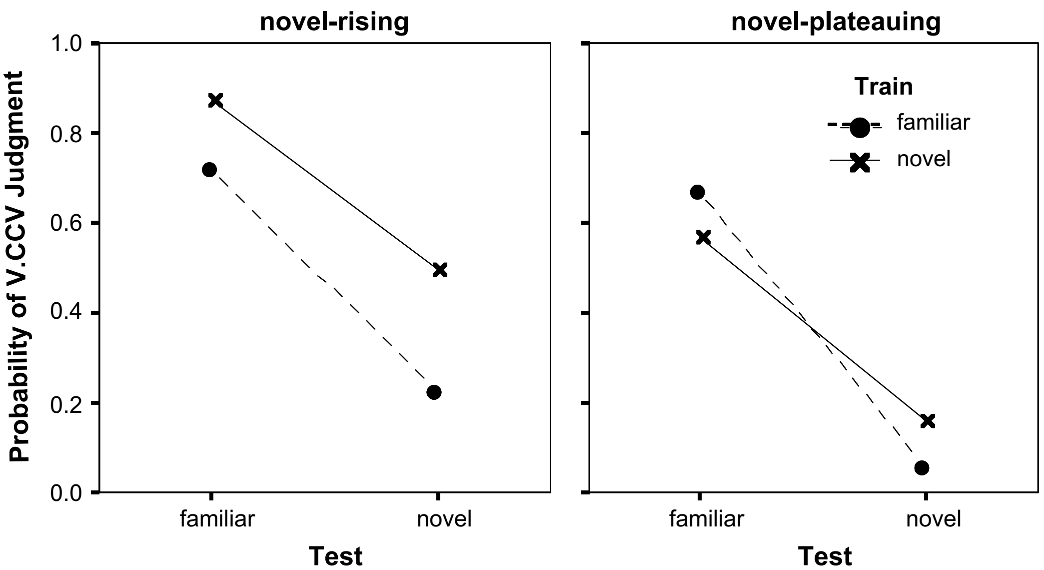

Fig. 2. The predicted probability of onset cluster judgments on word medial consonant sequences is graphed by sequence type, training, and test type. The within group differences at test are connected by lines. The between group differences are shown in separate panels (novel sequence type) and with distinct symbols (familiar or novel word onset sequences during training).

sequences may have significantly expanded their inventory of possible onset clusters, making them more likely to hear all medial sequences as onset clusters. Such an interpretation is consistent with results from Schiller, Meyer, and Levelt (1997) showing that set composition influences syllabification behavior.

Table 2 indicates that the effect of training was robust across nearly every novel-rising sequence. In many cases, the increase in onset cluster syllabification due to training was substantial: onset cluster judgments more than doubled in five out of eight novel-rising sequences. This kind of substantial increase also occurred for two out of eight novel-plateauing sequences, presumably leading to the significant training × test type interaction within the novel-plateauing condition. It is clear, though, from both Fig. 2 and Table 2 that novel-rising sequences were more readily treated as onset clusters than novel-plateauing sequences. This result is consistent with the expectation that listeners would extract novel onset phonotactics more readily from onset sequences that exhibited a greater degree of coarticulation and overlap with the following vowel than from those that were not obviously grouped in this way.

2.2.2. Explaining onset cluster judgments

Although the match between the acoustic and learning results are consistent with the expectation that more tightly coarticulated sequences are more often treated as onset clusters than less tightly coarticulated sequences, an analysis that used the C1:C2 duration measure to predict syllabification judgments showed that the afore-

Table 2
Syllabification judgments are displayed by training condition and novel sequence type

Novel-rising	Control (train: familiar)				Experimental (train: novel)			
	V.CCV	VC.CV	Other	*N*	V.CCV	VC.CV	Other	*N*
bn	6.8	88.6	4.5	44	39.1	45.7	15.2	46
dl	23.3	69.8	7.0	43	66.7	27.1	6.3	48
dn	21.2	57.6	21.2	33	22.2	38.9	38.9	36
gn	15.9	79.5	4.5	44	46.8	48.9	4.3	47
tl	19.0	64.3	16.7	42	43.2	34.1	22.7	44
vl	39.5	51.2	9.3	43	59.6	38.3	2.1	47
zr	27.3	59.1	13.6	44	33.3	27.1	39.6	48
žn	14.0	76.7	9.3	43	39.6	50.0	10.4	48
Novel-plateauing								
bd	6.8	50.0	43.2	44	7.5	32.5	60.0	40
dv	2.3	59.1	38.6	44	0.0	67.5	32.5	40
mn	4.7	46.5	48.8	43	10.3	27.5	62.2	40
mr	6.8	56.8	36.4	44	25.0	45.0	30.0	40
tk	0.0	72.7	27.3	44	0.0	62.5	37.5	40
vs	6.8	77.3	15.9	44	5.3	47.4	47.3	38
zv	6.8	79.5	13.6	44	10.0	70.0	30.0	40
žv	4.7	58.1	37.2	43	7.5	57.5	35.0	40

The table only shows syllabification judgments for medial novel sequences of both types and excludes those that were made on familiar sequences.

mentioned match between acoustics and differential learning may be qualitative rather than quantitative. Specifically, a random effects binary logistic regression analysis modeling the experimental groups' judgments indicated that the C1:C2 duration ratio interacted with sequence and test type [$\chi^2(3,\ N = 1133) = 9.31$, $p = .026$], such that the long–short pattern predicted listener onset cluster judgments on novel-plateauing sequences [$\chi^2(1,\ N = 1133) = 7.73$, $p = .003$], but not on novel-rising sequences [$\chi^2(1,\ N = 1133) = 0.87$, $p = .350$] or on familiar controls in the novel-rising or -plateauing conditions [rising, $\chi^2(1,\ N = 1133) = 0.02$, $p = .875$; plateauing, $\chi^2(1,\ N = 1133) = 0.99$, $p = .321$].

One possible explanation for these results is that all familiar and novel-rising word onset sequences were perceived as having reached some minimal threshold of coarticulation, and so further differences in the degree of coarticulation did not affect how often these were treated as onset clusters to the second syllable when the sequences appeared in word medial position. An alternative possibility is that the markedness effect on phonotactic learning has nothing to do with coarticulation, and instead can be attributed to other perceptual or cognitive factors.

An alternative perceptual–cognitive explanation for the results is suggested by the literature on the misapprehensions of phonotactically illegal sequences (e.g., Berent, Steriade, Lennertz, & Vaknin, 2007; Massaro & Cohen, 1983; Pitt, 1998). It could be that novel-rising sequences were sufficiently similar to familiar onset sequences and so were incorrectly mapped to legal familiar onset clusters before being syllabified as onset clusters. This explanation for the results correctly predicts the higher numbers of onset cluster syllabifications on novel-rising sequences, but does not explain the significant interactions attained between training and test, which indicated learning. Still, the possibility that misapprehension affected phonotactic learning is consistent with the fact that all of the 'other' judgments on novel-rising sequences, and 73% of those on the novel-plateauing sequence, were consonant–vowel–consonant judgments or singleton judgments, and both types of judgments are misperceptions. Further, these 'other' judgments were made more frequently on novel sequences than on familiar sequences [familiar vs. rising, $B = -0.64$, $t(2713) = -.3.63$, $p < .001$; familiar vs. plateauing, $B = -1.59$, $t(2713) = -9.89$, $p < .001$].

Given the higher number of obvious misperceptions on novel sequences, individual onset cluster responses were reexamined for phoneme misperceptions to assess the effect of misapprehension on learning. In the critical case of exposure to novel word onsets and subsequent onset cluster judgments on the same sequences in medial position, the results showed that 13.66% of the novel-rising and 46.16% of the novel-plateauing sequences had been misperceived, but that less than half of these misperceptions resulted in legal onset clusters. Specifically, 3.11% of onset cluster judgments on novel-rising medial sequences were misapprehended as legal onsets, and 19.23% of those on novel-plateauing sequences were misapprehended as such. The results therefore undercut the possibility that misapprehension accounted for the training-dependent increases in V.CCV judgments in the novel-rising experimental condition, though some increase in onset cluster syllabifications of novel-plateauing sequences may well have been due to misapprehension.

3. Experiment 2: Simulated consonant cohesion and phonotactic learning

The results presented so far suggest that phonotactic learning incorporates both information about consonant co-occurrences and information about how consonant sequences are produced. Consonant sequences that are more coarticulated are more likely to be treated as onset clusters, provided that listeners have learned that the sequences are also legal in word onset position. But how exactly does coarticulation affect phonotactic learning?

If we continue to assume that coarticulation is inferred from signal acoustics, one possibility is that the degree to which consonants are coarticulated or not interferes with the initial representation of novel phonotactics: word onset sequences that are perceived to be minimally coarticulated may be learned as simple sound sequences rather than as a structural unit (i.e., an onset cluster). Another possibility is that coarticulation interferes with the generalization of phonotactic knowledge: word medial sequences that are perceived as minimally coarticulated are assumed to belong to separate production units and so cannot be syllabified as onset clusters to the second syllable regardless of phonotactic legality. These possibilities were controlled in Experiment 2 in order to test the hypothesis that the native Russian productions of the consonant sequences constrained phonotactic learning in such a way as to produce more onset cluster syllabifications of novel-rising sequences than of novel-plateauing sequences.

Experiment 2 replicated in full the phonotactic learning experiment reported in Experiment 1, but the natural stimuli were edited to simulate greater and more consistent consonant–consonant coarticulation for all stimulus types in word onset and word medial position. In particular, all sequences were given durational patterns that matched the average durational pattern of the familiar-rising sequences. The edited stimuli were therefore expected to induce listeners to treat novel-rising and novel-plateauing sequences in the same way, and to induce them to make more onset cluster judgments overall.

3.1. Method

3.1.1. Participants

Forty-six new University of Oregon students participated in the phonotactic learning experiment in exchange for course credit. As before, all were native American-English speakers and all reported normal hearing.

3.1.2. Stimuli

All CVCVC stimuli from Experiment 1 were used in Experiment 2. All CCVCVC and CVCCVC stimuli were edited to create consistently high C1:C2 duration ratios in word onset and word medial position. A target C1:C2 duration ratio was calculated per speaker based on the average C1:C2 duration ratio produced by that speaker on familiar sonority-rising sequences in medial position. The consonant sequences were then edited so that they approximated the target. The procedure was (1) to increase C1 duration by copying and inserting a portion of the steady state energy associated with C1 closure, and (2) to reduce C2 duration by excising a portion of the steady state energy associated with C2 closure. The overriding aim, though, was to create natural sounding stimuli with different and more consistently high C1:C2 ratios. Given this aim, editing followed from the characteristic tempo of individual stimulus items rather than from a strict formula of percent excision/insertion. The amount of steady state energy that was excised or inserted per stimulus item was further limited by the subjective naturalness and perceptual saliency of the resulting sound wave. Fig. 3 shows the results of the editing procedure.

In spite of the subjective naturalness criteria, Fig. 3 shows that editing resulted in mean C1:C2 duration ratios that were much more consistent across word position and sequence types. Importantly, novel-rising and -plateauing sequences were given C1:C2 duration profiles that were more similar to the familiar sequences than to the natural production of these sequences by Russian speakers. Still, the C1:C2 duration ratios varied with sequence type [$F(2,30) = 4.36$, $p = .022$] and word position [$F(1,31) = 4.31$, $p = .046$], albeit much less than in the natural tokens (see the acoustic results reported in Section 2.1). The interaction between type and word position was not significant.

3.1.3. Procedures

The procedures were exactly the same as in Experiment 1. One control group ($N = 11$) was exposed to the familiar word onset sequences during training and to

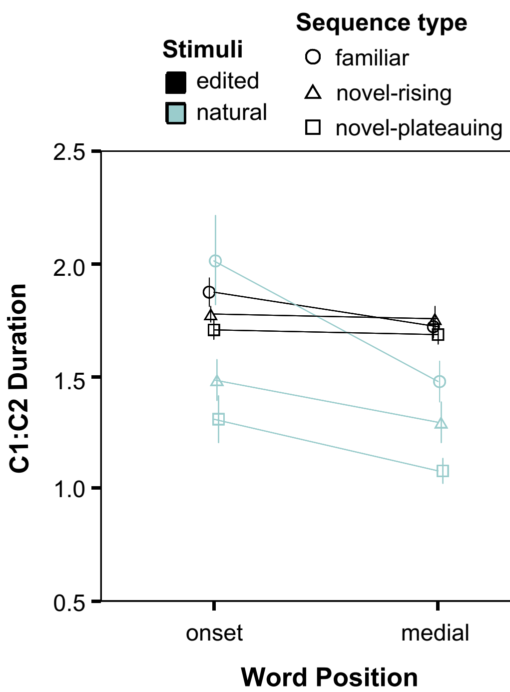

Fig. 3. The edited and natural consonantal duration ratios are shown for the Russian nonce word stimuli. Experiment 1 used the natural productions. Experiment 2 used edited tokens to control for C1:C2 effects on phonotactic learning.

familiar and novel-rising medial sequences during test. The matched experimental group ($N = 11$) was exposed to the novel-rising word onset sequences during training and the same mix of familiar and novel sonority-rising medial sequences during test. The other control group ($N = 12$) was exposed to familiar word onset sequences during training and to familiar and novel-plateauing medial sequences during test. The matched experimental group ($N = 12$) was exposed to novel-plateauing word onset sequences during training and the same mix of familiar and novel-plateauing medial sequences during test.

3.1.4. Response coding

As in Experiment 1, listener responses to the stimuli were coded as categorical syllabification judgments. A random effects logistic regression analysis predicted onset cluster syllabification (1 = V.CCV; 0 = VC.CV or VCC.V) according to test type (familiar or novel), which was embedded within listener (the random factor), listener was in turn embedded within training (familiar or novel), and training was embedded within sequence type (novel-rising or novel-plateauing).

3.2. Results and discussion

The results from Experiment 2 were largely the same as those from Experiment 1 contra the expectation that more consistent durational patterns across stimulus types would lead listeners (1) to treat novel-rising and novel-plateauing sequences in the same way, and (2) to make more onset cluster judgments overall. The only major difference between the two experiments was that the markedness effect on learning was clearer in Experiment 2 than in Experiment 1. A training dependent increase in onset cluster syllabification was only observed in the novel-rising condition. No learning effects were observed in the novel-plateauing condition. These and other results are reported in more detail below.

Once again, the categorical V.CCV, VC.CV, and VCC.V judgments were analyzed in a random effects logistic regression model. The overall fit of the model was good. The generalized χ^2 was 2216.07, yielding a generalized χ^2 to degrees of freedom ratio of 0.98. The effects of individual predictor variables were similar to those described in Experiment 1, with a few notable differences. As in Experiment 1, familiar medial sequences were syllabified as onset clusters to the second syllable more frequently than novel medial sequences whether the novel sequences were rising [$B = 2.83$, $t(2227) = 13.77$, $p < .001$] or plateauing [$B = 3.11$, $t(2227) = 14.48$, $p < .001$]. In contrast, training had no effect on the overall character of judgments in the two sequence conditions [rising, $B = -0.68$, $t(2227) = -0.80$, $p = .424$; plateauing, $B = 0.27$, $t(2227) = 0.34$, $p = .735$], and only interacted significantly with test type in the novel-rising condition [rising, $B = 2.83$, $t(2227) = 13.77$, $p < .001$; plateauing, $B = 0.20$, $t(2227) = 0.91$, $p = .362$]. Not surprisingly, these results did not change when the edited C1:C2 durations were added to the model.

Fig. 4 shows a clear markedness effect on phonotactic learning. Listeners exposed to novel onset sequences during training were more likely to syllabify these same sequences as onset clusters to the second syllable in word medial position, but those exposed to

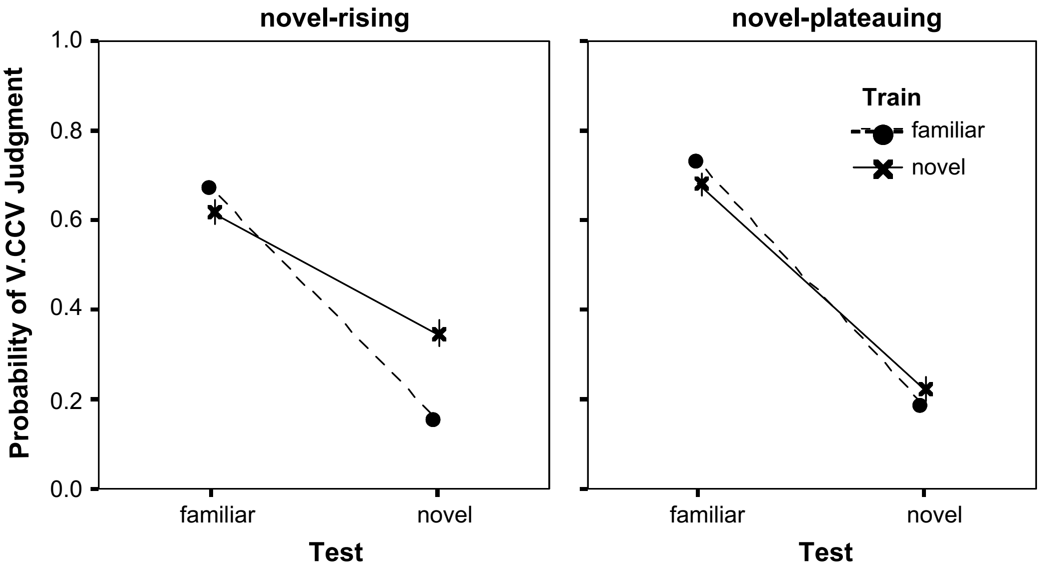

Fig. 4. The predicted probability of onset cluster judgments on word medial consonant sequences is graphed by sequence type, training, and test type.

novel-plateauing sequences during training were no more likely than the control group to syllabify the novel sequences as onset clusters to the second syllable at test.

Overall, the results suggest that listeners are no more or less likely to learn novel phonotactics from word initial sequences when an acoustic correlate of consonant–consonant coarticulation is controlled than when it is not. Even when all sequence types in both word onset and word medial position were edited to obtain roughly the same degree of C1 to C2 duration, the markedness effect on learning persisted. This result was surprising given the qualitative match between the consonantal duration patterns and onset cluster syllabification in Experiment 1, and the fact that the edited novel-plateauing sequences had C1:C2 durations that were even higher than those associated with naturally produced novel-rising sequences (see Fig. 3). The current results are also at odds with the quantitative result from Experiment 1, which showed that C1:C2 duration ratios in naturally produced stimuli predicted syllabification judgments on the novel-plateauing medial sequences.

4. Experiment 3: Self-productions and phonotactic learning

Both the qualitative and quantitative match between signal acoustics and listener judgments in Experiment 1 suggested that coarticulation constrains phonotactic learning in such a way as to give rise to a markedness effect. However, the results from Experiment 2 indicated that the markedness effect on learning emerges even when coarticulatory information in the signal is controlled. How might we reconcile these contradictory results? One way is to posit that the coarticulatory constraint on phonotactic learning is endogenous. Specifically, listeners may reinterpret what they hear according to their own production routines to encode the representations from which they abstract novel phonotactics. Assuming a universal (default) chunking algorithm, it is possible that the English listeners in Experiment 2 reconstituted the coarticulatory patterns observed in the natural Russian productions of sonority-rising and sonority-plateauing sequences by simulating their production during learning. Internally generated patterns of coarticulation could explain why listeners persisted in treating novel-rising and novel-plateauing sequences differently when C1:C2 durations were controlled in the input. Experiment 3 tested this possibility.

The train-test paradigm was used once again in Experiment 3, but this time all listeners were exposed to novel phonotactics during training; that is, to stimuli with novel-rising and novel-plateauing word onset sequences. During test, the listeners heard both the familiar and novel sequences in word medial position, just as in Experiments 1 and 2. In this way, Experiment 3 focused on group differences due to sequence type and ignored learning effects due to training condition (i.e., familiar vs. novel word onsets).

Only the edited stimuli from Experiment 2 were used in Experiment 3 to control for the C1:C2 duration differences present in the naturally produced stimuli. Listeners completed the training portion of the experiment as before, but during the test they were asked to repeat the stimulus item before providing a written syllabification of that item. Listener productions were recorded and medial C1 and C2 durations were measured. The C1:C2 duration ratios were then compared to the ratios observed in the naturally produced Russian stimuli and to the ratios obtained for the edited stimuli. The hypothesis that sonority-rising sequences are naturally more coarticulated than sonority-plateauing sequences predicts that the native English productions of the different sequence types will have C1:C2 duration ratios more similar to those of the native Russian productions than to those of the edited stimuli. The listeners' C1:C2 durations were also used to predict medial syllabifications of novel-rising and novel-plateauing sequences. The hypothesis that coarticulation constrains phonotactic learning predicts that the C1:C2 duration ratios associated with self-productions will explain listener syllabification judgments for all novel sequence types.

4.1. Method

4.1.1. Participants

Eighteen new undergraduate students from the University of Oregon participated in the phonotactic learning and production experiment in exchange for course credit. As before, all were native American-English speakers and all reported normal hearing.

4.1.2. Stimuli

The edited stimuli from Experiment 2 were used in Experiment 3.

4.1.3. Procedure

The procedures deviated in two ways from those used in the other phonotactic learning experiments. First, participants were only ever trained on words with novel onset sequences. Participants were exposed either to sonority-rising onset sequences ($N = 9$) or sonority-plateauing onset sequences ($N = 9$). Control groups who heard only familiar word onset sequences were not used in Experiment 3 because learning effects due to training on novel word onsets were documented in Experiment 1 and replicated in Experiment 2.

Second, participants were seated in a sound attenuated booth in front of a microphone and asked to repeat the stimulus item during test before providing a written syllabification of the word. Participants were coached to produce only fluent speech, and to only speak the stimulus item itself (i.e., not the frame sentence with the item). Only one participant thought that the instruction was to segment the stimulus items aloud, and this participant was quickly corrected before proceeding with the experiment. Productions were recorded using a high-quality microphone and a flash recorder. The digital files were later transferred to a desktop computer for acoustic analysis.

Participants had remarkably little trouble with the task. Some attempted Russian pronunciation (e.g., trilled *r*), but many others anglicized the stimuli. All participants reproduced the same iambic stress pattern that the Russian speakers had used when the stimuli were first recorded.

4.1.4. Measures

Durational measures were made just as in Experiment 1. The same segmentation criteria were applied. If listeners produced the American-English bunched *r* and this resulted in ambiguity in boundary placement, the ambiguity was resolved by following the excursion of the 3rd formant into the following vowel and by placing the boundary at the temporal midpoint of the transition (Klatt, 1976).

As before, the consonantal duration pattern was captured by the C1:C2 duration ratio. Duration ratios were not calculated on the tokens that were produced with an epenthetic vowel dividing the medial consonant sequence. Ratios were also not obtained on tokens where the participant produced only 1 of the 2 consonants in medial position. The ratios obtained on the remaining 82.8% of productions ($N = 946$) were compared across sequence type and to those obtained from the four native Russian speakers' productions (i.e., the stimuli used in Experiment 1, $N = 96$).

4.1.5. Response coding

As in Experiments 1 and 2, written syllabifications were coded as categorical boundary judgments. Random effects logistic regression analyses were used to predict onset cluster syllabification (1 = V.CCV; 0 = VC.CV or VCC.V) based on exposure to a particular sequence type (novel-rising or novel-plateauing), test type (familiar or novel), and C1:C2 duration. Both C1:C2 duration and test type were embedded within listener, the random effect. Listeners were embedded within sequence type.

4.2. Results and discussion

The acoustic results indicated that English listeners produce novel-rising and novel-plateauing medial sequences with the same consonantal duration pattern as the Russian speakers, but that they produced familiar consonant sequences differently and with C1:C2 duration ratios that were even higher than those of the edited stimuli. The analyses on medial syllabification judgments showed that listeners' self-productions affected their syllabification behavior: sequences produced with higher C1:C2 duration ratios were more likely to be syllabified as onset clusters to the following vowel than sequences produced with lower C1:C2 duration ratios. However, the analyses also showed that the effect of self-production varied with sequence type. In general, listeners' self-productions predicted their behavior on novel sequences better than on familiar sequences. Listeners were also somewhat more inclined to syllabify novel-rising sequences as onset clusters than novel-plateauing sequences given the same C1:C2 duration ratios. These overall results are presented in more detail below.

4.2.1. Consonantal duration patterns in listeners' self-productions

The first analysis examined whether the native English participants produced different C1:C2 duration ratios for familiar and novel sequences, and whether novel sequence type also affected C1:C2 duration. The results showed that familiar medial sequences had much higher C1:C2 duration ratios than novel medial sequences [$F(1, 927) = 152.25$, $p < .001$], and that there was no overall effect of novel sequence type (rising vs. plateauing) on C1:C2 duration. A post hoc comparison of the medial novel-rising and novel-plateauing productions also showed no significant difference in the C1:C2 duration values associated with just these sequences [$B = -0.05$, $t(927) = -0.36$, $p = .721$]. These results are evident in Fig. 5, which shows the esti-

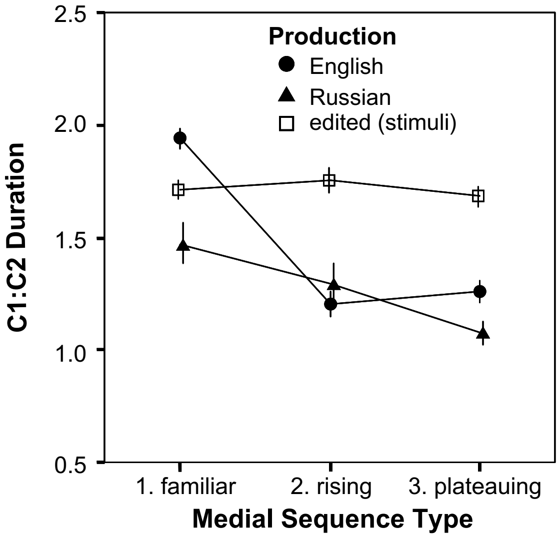

Fig. 5. The consonantal duration ratios are shown for the English and Russian productions of the different nonce word stimuli. Native English-speaking participants were presented with the edited versions of the stimuli to repeat. The edited consonantal duration ratios are shown with the natural Russian values for comparison to the English values.

mated marginal C1:C2 duration means for the different medial sequences produced by the English listeners.

To facilitate a comparison with the naturally produced and edited stimuli, Fig. 5 also shows the estimated marginal mean C1:C2 durations for the Russian productions of familiar, novel-rising, and novel-plateauing sequences as well as the marginal mean values associated with the edited sequences. The marginal mean C1:C2 duration ratios for English and Russian productions of familiar sequences were 1.94 (.04) and 1.48 (.17) respectively. Those for novel-rising sequences were 1.20 (.05) and 1.29 (.14) for English and Russian, and those for novel-plateauing sequences were 1.26 (.05) and 1.08 (.12). Only the mean difference between the English and Russian productions of the familiar sequence was significant after adjustment for multiple comparisons [mean difference = .47, $SE = .18$, $p < .05$]. The English and Russian productions of novel-rising and -plateauing sequences were not statistically different from one another.

In contrast to these results, neither the English nor Russian mean C1:C2 durations associated with familiar sequences were statistically different from the edited mean of 1.72 (.17); however, the mean differences between English and the edited C1:C2 durations and between Russian and the edited values were significantly different for both the novel-rising sequences [Eng. vs. edited, mean difference = .55, $SE = .14$, $p < .01$; Ru. vs. edited, mean difference = .46, $SE = .19$, $p < .05$] and for the novel-plateauing sequences [Eng. vs. edited, mean difference = .43, $SE = .12$, $p < .01$; Ru. vs. edited, mean difference = .61, $SE = .19$, $p < .05$].

Overall, the comparison between English, Russian, and the edited C1:C2 duration ratios confirms the prediction that the English-speaking participants would produce the stimuli in a manner more consistent with the natural Russian productions than with the edited productions.

4.2.2. Effect of self-productions on onset cluster judgments

The next analysis sought to determine whether the participants' productions would predict their syllabifications of medial familiar and novel-rising and -plateauing sequences in much the same way as the natural Russian C1:C2 durations predicted judgments on novel-plateauing sequences in Experiment 1.

Once again, the categorical V.CCV, VC.CV, and VCC.V judgments were analyzed in a random effects logistic regression model, but with C1:C2 duration included as a predictor variable embedded within listener. The model provided a good fit of the data. The generalized χ^2 was 786.34, yielding a generalized χ^2 to degrees of freedom ratio of 0.92. A look at the predictor variables showed that C1:C2 duration was a strong overall predictor of syllabification [$\chi^2(1, N = 842) = 22.57$, $p < .001$], albeit one that interacted with sequence and test type [$\chi^2(3, N = 842) = 15.96$, $p = .001$]. This interaction is shown in Fig. 6.

It is evident from Fig. 6 that consonant–consonant coarticulation, as measured by the C1:C2 duration ratio, had a smaller effect on syllabification of familiar medial sequences than it did on the syllabification of novel medial sequences. Familiar sequences were usually treated as onset clusters, and higher C1:C2 durations made this outcome only slightly more likely than lower C1:C2 durations [rising(familiar)

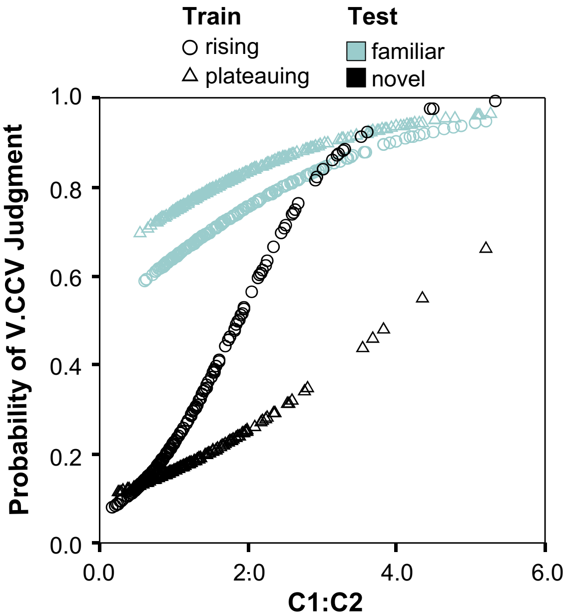

Fig. 6. The probability of onset cluster judgments on word medial consonant sequences as predicted by the medial consonantal duration ratio (self-productions) and by test type (English familiar or novel syllable–onset sequences).

$B = 0.34$, $t(842) = 1.56$, $p = .120$; plateauing(familiar) $B = 0.02$, $t(842) = 0.12$, $p = .906$]. In contrast, novel-rising and -plateauing sequences were significantly more likely to be treated as onset clusters when these were produced with higher C1:C2 duration ratios than when they were produced with lower ratios [rising(novel) $B = 1.43$, $t(842) = 4.31$, $p < .001$; plateauing(novel) $B = 0.95$, $t(842) = 2.54$, $p = .011$]. The results displayed in Fig. 6 also suggest that participants were more likely to treat novel-rising sequences as onset clusters than novel-plateauing sequences given the same C1:C2 duration ratio. The C1:C2 slope associated with judgments on novel-rising sequences was steeper than the slope associated with judgments on novel-plateauing, though the difference between the slopes was not significant [$B = 0.48$, $t(842) = 0.95$, $p = .343$].

Overall, these results suggest that listeners reference their own productions to assess the degree to which consonants are coarticulated and overlapped with a subsequent vowel, and that the extent to which novel sequences are coarticulated in this way influences how listeners generalize novel word onsets to syllable onsets. The results also indicate that familiar sequences are syllabified largely without reference to self-productions, and that the markedness effect on phonotactic learning may persists somewhat even when self-produced consonantal duration patterns are controlled. These latter results might be explained if we assume that phonological knowledge is ultimately abstracted away from the motoric and acoustic detail provided by simulated speech. Syllabification of familiar onset sequences may proceed entirely from abstracted knowledge and without reference to self-productions. In contrast, exposure to novel word onset phonotactics may lead to initial abstract representations of novel syllable onsets, but medial syllabification of these same sequences still requires additional detailed phonetic information, which is provided through simulation.

5. General discussion

To summarize, acoustic measures of the stimuli in Experiment 1 indicated that sonority-rising sequences are more likely to be coarticulated than sonority-plateauing sequences, as measured by the ratio of C1 to C2 duration – an acoustic correlate of coarticulation and well studied cue to boundary location. Results from the phonotactic learning experiment showed that listeners more readily learned novel onset phonotactics from the sonority-rising word onset sequences than from sonority-plateauing sequences. The match between signal acoustics and listener judgments suggested that coarticulation may constrain learning in a manner consistent with the SSP. Results from Experiment 2 indicated that, if this constraint is real, it is likely endogenous to the learner: a clear markedness effect on phonotactic learning was observed even though the acoustics of the signal were modified to control for differences in C1 to C2 duration. Results from Experiment 3 were consistent with an endogenous production constraint on phonotactic learning. English-speaking listeners reconstituted the different consonantal duration patterns associated with Russian productions of familiar and novel medial sequences. More importantly, listeners' self-productions predicted their syllabification judgments on novel word medial sequences. Together, the results suggest that the formation of new onset cluster representations is influenced by the production of consonant sequences, which is similar in English and Russian.

Overall, the results indicate that a markedness constraint on phonotactic learning may arise from the default organization of sequential articulatory action. Insofar as the markedness effect parallels predictions made from the typological facts, the results imply that the typological facts may be explained with reference to a universal motor routine rather than with reference to cognitive principles specific to language. In particular, the suggestion here is that the typologically frequent sequencing pattern characterized by the SSP emerges from a default, perceptually honed motor routine that groups segments within a single open–close cycle of the jaw. The link between coarticulatory grouping and syllables means that this default routine also provides a foundation for syllabic representations and so constrains phonotactic learning directly.

The hypothesis that basic motor processes constrain on-line learning implies that phonetics can actively shape the phonology. This implication obviously conflicts with the classical generativist view that markedness principles represent innately specified, abstract linguistic knowledge. However, it also at odds with a view advocated by some who otherwise embrace the idea that phonetics shapes phonology (e.g., Flemming, 2001; Hayes & Steriade, 2004; Wilson, 2006). To illuminate the difference between the view adopted here and an important alternative view on the relationship between phonetics and phonology, let me expand on the current proposal.

The proposal is that motor processes constrain on-line learning by providing structured input for phonological abstraction. Consonant sequences that are easily coarticulated are available to be abstracted as syllable-onset clusters. Those sequences that are not so easily coarticulated are not available for such abstraction. The phonetics therefore constrains the phonology by filtering the input. One implication of this view is that phonetics interacts with, but is still independent from, the phonology. A further implication is that the phonology only encodes possible, or productive, sound patterns; it does not explicitly restrict impossible patterns or constrain output. The assumption is that such restrictions are already implicit in the more concrete representations and overall structure of the lexicon.

Both implications are at odds with an alternative view that collapses phonetics into phonology (e.g., Flemming, 2001; Hayes & Steriade, 2004; Wilson, 2006). According to this alternative, phonetic constraints are directly encoded in the phonology. For example, Hayes and Steriade (2004) hypothesize that universal markedness constraints are emergent from the phonetic knowledge that an infant acquires through babbling and other vocal–auditory experiences. When the phonetically grounded phonological constraints are acquired, all linguistically relevant information on possible and impossible patterns is encoded in the phonology. Phonetics ceases to affect the grammar. Instead, phonetics comes to merely reflect phonological constraints on output. That is, the relationship between phonetics and phonology becomes uni-directional in the mature language user: phonology impacts on-line phonetics, but the reverse is not true. Such a conclusion appears to be inconsistent with the present findings, which instead suggest that the variable way in which one produces an onset sequence helps to determine the types of abstractions that are made.

Although the present universalist explanation for a markedness effect on phonotactic learning may differ from both the classical generativist explanation and from alternative phonetically based phonology explanations, it shares with these explanations the assumption that the sound patterns of language are constrained by our biology. Still, however likely it is that all humans organize sequential articulatory action according to some default motor routine, it is equally clear that we all acquire language-specific routines that may influence subsequent learning. Consider, for example, the difficulty we have in producing (and so acquiring) the sounds and sound patterns of a second language. In the context of the present study, which used naturally produced Russian nonwords as stimuli to assess English-speaking listeners' ability to extract novel phonotactics, it may be that a language-specific explanation of the results is more appropriate than a universalist explanation. Experiment 3 provides a case in point.

Although the native English speakers' productions in Experiment 3 were similar to the natural Russian productions of the different consonant sequences, there were also clear differences between the groups. In particular, native English speakers' productions of familiar sequences yielded much larger C1:C2 duration ratios on average than Russian productions of the same sequences. This difference is not surprising given previous work on the relative durations of different consonant types in English and Russian. Recall that the familiar sequences were all stop-liquid sequences. In

Redford (2007), I measured stop-liquid sequences in English and Russian (and Finnish) in word-internal and word-edge position to differentiate between prosodic and articulatory effects on consonantal duration patterns. The results that are relevant to the current discussion were that English speakers reduced post-consonantal, pre-vocalic liquids more than Russian speakers, and they produced voiceless stops with more aspiration than Russian speakers, except when these occurred in word-final position. Such results suggest that, compared to Russian speakers, English speakers normally produce especially long stops (closure + release) and especially short liquids when these consonants are uninterrupted by a word boundary and occur in sequence, as they did in the present study.

If language experience explains the difference between English and Russian speakers' production of familiar sequences, perhaps language experience can also explain the difference between English speakers' production of familiar and novel sequences. While English speakers produced familiar sequences with high C1:C2 durations, they produced novel-rising and novel-plateauing sequences with lower C1:C2 durations. Moreover, the novel-rising and novel-plateauing sequences were produced (on average) with the same C1:C2 duration. Such results are not consistent with a default production routine, which is hypothesized to group rising sequences more readily than plateauing sequences. However, the results are consistent with work on English speakers' productions of novel consonant clusters. Davidson (2006b) has shown that English speakers do not apply native language patterns of consonant cluster coordination to nonnative onset sequences. Instead, novel sequences are inappropriately 'pulled apart' (less coarticulated) during production than familiar sequences. This means that English speakers in the present study might not have coarticulated novel sequences as tightly as familiar sequences simply because the sequences were novel; not because they differed in the extent to which they could be coarticulated according to some default motor routine. On the other hand, Experiment 1 showed that Russian speakers also coarticulate familiar sequences more tightly than novel-rising and novel-plateauing sequences (see Fig. 1). This suggests either that the similarities between native English and native Russian productions were fortitous or that English and Russian speakers have different language-specific motor routines, but nonetheless also share a basic mode for organizing sequential articulatory action.

Whereas the current results may be ambiguous with respect to the hypothesis of a shared routine for organizing sequential action in speech, the results unambiguously support the hypothesis that production constrains phonotactic learning. The more tightly a novel sequence was coarticulated and overlapped with the following vowel, the more likely it was to be syllabified as an onset cluster to the following syllable. The fact that this result was based on self-productions strongly suggests that coarticulation is inferred from simulating the production of these sequences rather than from the acoustic correlates of coarticulation that are perceptually available in the signal. This is not to suggest that phonological behavior always depends on the facts of implementation. After all, previous work on syllabification suggests a myriad of influences on listeners' judgments apart from the subphonemic acoustic patterns focused in the present study (Content, Kearns, & Frauenfelder, 2001; Redford & Randall, 2005; Schiller et al., 1997; Smith & Pitt, 1999; Treiman & Zukowski, 1990). Rather, the suggestion is that early representations are strongly tied to the facts of production, and perhaps to little else, in a way that later representations are not.

More generally, the present results suggest that the powerful effects of statistical learning may be constrained by the perception–production loop in language. This loop, while universally recognized, is not typically emphasized in discussions of linguistic representation. Instead, when phonetics is invoked to explain universal language structure, the emphasis is usually on perceptual processes (e.g. Newport & Aslin, 2004; Wilson, 2006; see also Blevins, 2004; Hayes et al., 2004; Hume & Johnson, 2001). Although such processes might help account for markedness constraints on phonological learning, they do not account for why language users have relatively stable linguistic representations given the highly variable input provided by everyday language encounters and the clear power of statistical learning.

Each individual has daily experience with multiple sources of language input that provide new information about language categories and overall language structure, but such input has little effect on linguistic behavior even though it is no doubt encoded as important language information. Consider, for example, the fact that exposure to child language, foreign accented language, or new dialects of the same language will expand an adult's ability to comprehend individuals speaking in any of these varieties, but will not typically affect the adult's output patterns or her intuitions about language structure. The perception–production loop provides an explanation for this stability in behavior. In particular, entrenchment in production may anchor lexical representations, which in turn could anchor phonological knowledge, insofar as such knowledge is emergent from the lexicon (Bybee, 2001; Pierrehumbert, 2001). This idea is akin to one proposed by Pierrehumbert (2001) in a discussion of historical lenition (vowel reduction, stop spirantization, final devoicing, and so on). Pierrehumbert suggests that category diffusion for a particular phoneme is constrained by practice effects. She also makes the larger point that psychological models, built to account for judgments on perceptual data, cannot account for language behavior unless they are amended to incorporate the effects of language production. Although Pierrehumbert's specific interest is in exemplar theory as a model of language behavior, her argument would also seem to apply to statistical learning theory, insofar as statistical learning of language patterns is about pattern extraction from perceptual input.

To conclude, statistical learning of novel language phonotactics is likely constrained by production; specifically, by the motor routine(s) that organize(s) sequential articulatory action for output. Such a proposal is consistent with a long tradition in phonetics that recognizes the role of both perception and production in shaping language sound structure (Demolin & Soquet, 1999; Lindblom et al., 1984; Martinet, 1955; Passy, 1890; Redford et al., 2001).

Acknowledgements

I am grateful to Tatiana Furrow, who helped to ensure that the stimuli were possible Russian words and correctly represented in the Cyrillic alphabet; to Robin High, resident statistical programmer and consultant at the University of Oregon, who consulted extensively with me on the statistics; and to the anonymous reviewers of *Cognition*, who provided valuable critical feedback on earlier drafts of this paper.

Appendix A

Target CCVCVC and CVCCVC stimuli written in normal English orthography

Onset	Medial
a. Nonce Russian words with familiar syllable-onset sequences	
blatat	tablat
bravat	vabrat
drevat	vadret
frenat	nafret
glatat	taglat
gramat	magrat
slatat	taslat
shragat	gashrat
b. Nonce Russian words with novel sonority-rising onset sequences	
bnagat	gabnat
dlatat	tadlat
dnevat	vadnet
gnavat	vagnat
tlevat	vatlet
vlatat	tavlat
zrenat	nazret
zhnamat	mazhnat
c. Nonce Russian words with novel sonority-plateauing onset sequences	
bdevat	vabdet
dvegat	gadvet
mnetat	tamnet
mrevat	vamret
tkamat	matkat
vsenat	navset
zvatat	tazvat
zhvasat	sazhvat

References

Anderson, S., & Port, R. (1994). Evidence for syllable structure, stress and juncture from segmental durations. *Journal of Phonetics, 22*, 283–315.

Bell, A., & Hooper, J. B. (1978). Issues and evidences in syllabic phonology. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 3–22). Amsterdam: North Holland.

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

Berg, T., & Niemi, J. (2000). Syllabification in Finnish and German: Onset filling versus onset maximization. *Journal of Phonetics, 28*, 187–216.

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

Boersma, P., Weenink, D. (2002). Praat: Doing phonetics by computer (version 4.0.34) [Computer program]. Retrieved November 2002 from http://www.praat.org/.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6*, 201–251.

Browman, C. P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics, 18*, 299–320.

Butt, M. (1992). Sonority and the explanation of syllable structure. *Linguistica Berichte, 137*, 45–67.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition, 87*, B69–B77.

Christie, W. M. (1977). Some multiple cues for juncture in English. *General Linguistics, 17*, 212–222.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*, 141–176.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Between the grammar and the physics of speech* (pp. 283–333). New York: CUP.

Content, A., Kearns, R. K., & Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language, 45*, 177–199.

Davidson, L. (2006a). Phonotactic and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science, 30*, 837–862.

Davidson, L. (2006b). Phonology, phonetics, or frequency: Influences on the production of nonnative sequences. *Journal of Phonetics, 34*, 104–137.

de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech, 36*, 197–212.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study on the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1355–1367.

Demolin, D., & Soquet, A. (1999). The role self-organization in the emergence of phonological systems. *Evolution of Communication, 3*, 21–48.

Erickson, D. (1998). Effects of contrastive emphasis on jaw opening. *Phonetica, 55*, 147–169.

Erickson, D., Fujimura, O., & Pardon, B. (1998). Articulatory correlates of emotion control: Emotion and emphasis. *Language and Speech, 41*, 399–417.

Fisher, C., Church, B. A., & Chambers, K. (2004). Learning to identify spoken words. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 3–40). Cambridge, MA: MIT Press.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology, 18*, 7–44.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at the edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728–3740.

Haggard, M. (1973). Correlations between successive segment durations: Values in clusters. *Journal of Phonetics, 1*, 111–116.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics, 23*, 305–322.

Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: An Acoustic Analysis of the Queen's Christmas Broadcasts. *Journal of the International Phonetic Association, 30*, 63–78.

Hayes, B., Kircher, R., & Steriade, D. (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hayes, B., & Steriade, D. (2004). Introduction: The phonetic basis of phonological markedness. In B. Hayes, R. Kircher, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 1–34). Cambridge: Cambridge University Press.

Hooper, J. B. (1972). The syllable in phonological theory. *Language, 48*, 525–540.

Hume, B., & Johnson, K. (2001). *The role of speech perception in phonology*. San Diego: Academic Press.

Kawasaki-Fukumori, H. (1992). An acoustical basis for universal phonotactic constraints. *Language and Speech, 35*, 73–86.

Kingston, J., & Diehl, R. (1994). Phonetic knowledge. *Language, 70*, 419–453.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208–1221.

Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics, 27*, 33–54.

Lindblom, B. (1983). Economy of speech gestures. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 217–245). New York: Springer-Verlag.

Lindblom, B., MacNeilage, P. F., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & O. Dahl (Eds.), *Explanations for Language Universals* (pp. 181–203). Berlin: Mouton.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499–546.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science, 288*, 527–531.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech. *Child Development, 71*, 153–163.

Martinet, A. (1955). *Economie des changement phonétiques*. Berne: Francke.

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

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*, 91–121.

Newport, E. L., & Aslin, R. N. (2000). Innately constrained learning: Blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual boston university conference on language development* (pp. 1–21). Sommerville, MA: Cascadilla.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127–162.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*, 21–46.

Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London: Longman.

Ohala, J. J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. Håkon Jahr (Eds.), *Language and its ecology: Essays in memory of Einar Haugen* (pp. 343–365). Berlin: Mouton de Gruyter.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition, 83*, B13–B23.

Passy, P. (1890). *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin, Didot.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006). The role of phonetic naturalness in phonological rule acquisition. In D. Bamman, R. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th annual boston university conference on language development* (pp. 464–475). Somerville, MA: Cascadilla Press.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46*, 115–154.

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

Pulgram, E. (1970). *Syllable, word, nexus, cursus*. The Hague: Mouton.

Redford, M. A. (2007). Word-internal versus word-peripheral consonantal duration patterns in three languages. *Journal of the Acoustical Society of America, 121*, 1665–1678.

Redford, M. A., Chen, C. C., & Miikkulainen, R. (2001). Constrained emergence of universals and variation in syllable systems. *Language and Speech, 44*, 27–56.

Redford, M. A., & Randall, P. (2005). The role of juncture cues and phonological knowledge in English syllabification judgments. *Journal of Phonetics, 33*, 27–46.

Rubach, J., & Booij, G. (1990). Syllable structure assignment in Polish. *Phonology, 7*, 121–158.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*, 110–114.

Schiller, N. O., Meyer, A. S., & Levelt, W. J. M. (1997). The syllabic structure of spoken words: Evidence from the syllabification of intervocalic consonants. *Language and Speech, 40*, 103–140.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development, 1*, 289–316.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (pp. 337–383). Dordrecht: Foris.

Smith, K. L., & Pitt, M. A. (1999). Phonological and morphological influences in the syllabification of spoken words. *Journal of Memory and Language, 41*, 199–222.

Steriade, D. (1999). Alternatives to syllable-based accounts of consonantal phonotactics. In O. Fujimura, B. Joseph, & B. Palek (Eds.), *Proceedings of the 1998 linguistics and phonetics conference* (pp. 205–242). The Karolinum Press.

Stone, M. (1981). Evidence for a rhythm pattern in speech production: Observations of jaw movement. *Journal of Phonetics, 9*, 109–120.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15*, 49–74.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27*, 87–104.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29*, 66–85.

Tuller, B., & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research, 34*, 501–508.

Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 387–398.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science, 30*, 945–982.

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kircher, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 34–57). Cambridge: Cambridge University Press.