Phonotactics and articulatory coordination interact in phonology:

evidence from non-native production

Lisa Davidson

RUNNING TITLE: Phonotactics and articulatory coordination interact in phonology

New York University
Department of Linguistics
719 Broadway, 4th Floor
New York, NY 10003
USA
phone: 212-992-8761
email: lisa.davidson@nyu.edu

Abstract

A core area of phonology is the study of phonotactics, or how sounds are linearly combined. Recent crosslinguistic analyses have shown that the phonology determines not only phonotactics, but also the articulatory coordination or timing of adjacent sounds. This paper explores how the relationship between coordination and phonotactics affects speakers producing non-native sequences. Recent experimental results (Davidson 2005, 2006) have shown that English speakers often repair unattested word-initial sequences (e.g. /zg/, /vz/) by producing the consonants with a less overlapping coordination. A theoretical account of the experimental results employs Gafos's (2002) constraint-based grammar of coordination. In addition to Gafos's ALIGNMENT constraints establishing temporal relationships between consonants, a family of RELEASE constraints is proposed to encode phonotactic restrictions. The interaction of ALIGNMENT and RELEASE constraints accounts for why speakers produce non-native sequences by failing to adequately overlap the articulation of the consonants. The Optimality Theoretic analysis also incorporates floating constraints to explain why speakers are not equally accurate on all unattested clusters.

1. Introduction

In the process of acquiring a second language, one phonological aspect of the target language that learners must discover is the phonotactics, or restrictions on how and where sounds may combine in linear sequences. For example, while /k/ and /t/ are phonemes of both English and Russian, only Russian allows the sounds to be combined word-initially, as in [kto] 'who'. Yet, the acquisition of phonotactics is not simply a matter of learning which consonants may be linearly arranged, but also how they are articulatorily coordinated, or timed, with respect to one another. For example, an English speaker acquiring Russian would easily learn that sequential stops are permissible word-medially, since she would have experience with words like *laptop* in English. However, the speaker would also have to learn that the first consonant in the sequence is fully released in Russian (i.e. the end of the constriction of the consonant is audible), whereas it is not released in English in the same position. This is illustrated in Fig. 1, showing spectrograms of *laptop* and Russian *vnaklatku* 'with sugar'. Zsiga (2003) has shown that English speakers learn the Russian word-medial pattern relatively quickly, but Russian speakers have more difficulty with the lack of release in English.

------------------------------Insert Figure 1 about here ----------------------------

A recent set of experimental studies has shown that English speakers producing Slavic consonant sequences in word-initial position display two interesting behaviors: (1) though all of the word-initial test sequences in the studies are unattested in English (e.g. /zg/, /fk/, /vn/, etc), speakers do not produce them with the same accuracy, and (2) when they fail to produce them accurately, it is often because they do not employ the correct coordination between the

consonants. The results of the first study (Davidson, 2006) indicate that English speakers producing fricative-initial sequences (such as in *zgano*) are more accurate on some of the non-native clusters than on others. Another result of the first study is that English speakers often produce the unattested consonant sequences with some vowel-like acoustic material present between the first and second consonants on the acoustic record. Yet, an examination of the acoustic characteristics of this transitional vowel demonstrate that it is not the same as a corresponding lexical (underlying) vowel. These differences suggest that the origin of the transitional vowel is something other than a phonemic schwa that has been epenthesized, or inserted by the phonological grammar, into the consonant cluster.

To further address this issue, the second study uses ultrasound imaging to confirm that the articulatory patterns of some speakers are not consistent with schwa epenthesis (Davidson, 2005). Instead, as will be described in detail in Section 3, the results of these studies indicate that English speakers are not sufficiently overlapping the articulations of the consonants in the cluster. That is, English speakers appear to be leaving too much time between the release of the constriction of the first consonant and the beginning of the constriction of the second consonant, giving rise to a vowel-like sound between the two consonants, such as /zgano/ → [z$^\text{ə}$gano]. In this notation, the superscript schwa represents the presence of schwa-like material between the two consonants that is not identical to English lexical schwa. A comparison of the pseudo-word *zgano* as produced by a Slovak speaker and by an English speaker attempting to reproduce the Slovak example is shown in Fig. 2. Fig. 2 also contains a spectrogram of the English speaker producing /zəgano/ with an intentional schwa between the initial consonants.

------------------------------Insert Figure 2 about here ----------------------------

The goal of this paper is to develop a theoretical account of how articulatory coordination and phonotactic restrictions interact as speakers attempt the production of non-native consonant sequences. Such an analysis requires three steps: (1) explaining why English speakers do not produce all non-native consonant clusters with equal accuracy, (2) providing evidence that when the sequences are repaired, it is by manipulating the articulatory coordination of the consonants, not by epenthesizing a vowel, and (3) developing a framework that can explain the preceding two observations by combining both phonotactic and coordination constraints in a single representationally coherent phonological system.

To accomplish the last of these steps, an account integrating elements of both Articulatory Phonology and Optimality Theory is proposed. Articulatory Phonology is a framework that places articulatory coordination between adjacent sounds into the phonological system of a language. The basic representational unit of Articulatory Phonology, called the gesture, is well-suited to the English production facts because it is discrete in that it is either present or absent in the output of the phonology, but continuous in that it has a lifespan and may be specified in time relative to other gestures at predefined intervals (Browman & Goldstein, 1986 et seq.). Optimality Theory (OT) is likewise appropriate, since its main strength is in demonstrating how multiple phonological pressures—here, phonotactics and coordination— compete with one another to determine the surface form (Prince & Smolensky, 1993/2004). Previous OT accounts of coordination have been proposed before (e.g., Gafos, 2002; Hall, 2003), but the interaction between constraints governing coordination and constraints pertaining to phonotactics has not yet been fully fleshed out.

The remainder of the paper is structured as follows. In Section 2, relevant concepts of Articulatory Phonology are briefly reviewed. In Section 3, experimental results from an acoustic

and articulatory study demonstrate how English speakers produce non-native word-initial consonant sequences. These data form the basis of the formal analysis in Section 4, which shows how coordination, phonetically-influenced phonotactics, and speaker variation are combined in an OT grammar. Section 5 contains concluding remarks.

2. Coordination in consonant sequences

Researchers developing the theoretical framework called Articulatory Phonology (AP) have argued that articulatory coordination among adjacent sounds, as in the cases described above, is a language-specific phenomenon that is controlled by phonological grammars. It has been demonstrated that there are articulatory differences in the production of the same sounds depending on their syllabic position (Browman & Goldstein, 1995; Byrd, 1996b; Kochetov, to appear; Krakow, 1999; Sproat & Fujimura, 1993), and it has also been shown that there are cross-linguistic differences in how sequential sounds are articulatorily coordinated (e.g. Bradley, 2002; Browman & Goldstein, 1990b; Byrd, 1995, 1996a; Gafos, 2002; Hall, 2003; Kochetov, 2001/2002; Zsiga, 2003).

The notion that the interconsonantal vowel-like sound produced by an English speaker attempting a word like *zgano* could emerge as a result of a particular coordination relationship between [z] and [g] is consistent with Gafos's (2002) account of transitional schwa in Moroccan Colloquial Arabic, and other languages such as Sierra Popoluca, Piro, and Hua (Elson, 1956; Hall, 2003; Matteson & Pike, 1958). Examples from these languages are given in (1).

(1)   Moroccan Arabic: [smimᵊn] 'fat, diminutive'
        Piro: [tᵊkatʃi] 'sun'
        Hua: [fᵊtu] 'smell'

Transitional vowels have previously been defined as (a) being in free variation, (b) having no syllabic stress, (c) being much shorter than all other vowels in the language, and (d) being heavily coarticulated with surrounding consonants.

Gafos (2002) develops a framework which formally captures the phonological nature of coordination relationships among different types of sequences. Following previous work in AP (Browman & Goldstein, 1986 et seq.), the *gesture* is taken to be the basic unit of phonological representation. A gesture differs from the traditional concept of phoneme in that it is defined for both spatial and temporal specifications. Spatial variables in AP include constriction location (e.g. place of articulation, such as coronal or labial) and constriction degree (e.g. manner of articulation, such as stop or fricative), and the temporal specification refers to how the gesture unfolds over time. A number of studies have shown that the lifespan of a gesture can be divided into a few discrete landmarks which are important for determining how gestures are coordinated with respect to one another (Browman & Goldstein, 1990a, 1990b, 1995; Byrd, 1995). Gestural landmarks are shown in Fig. 3 (slightly modified here).

-----------------------------Insert Figure 3 about here -------------------------------

In English, the first consonant in an obstruent (i.e. stop or fricative) cluster is usually not audibly released, as in the case of *laptop* in Fig. 1 (e.g. Henderson & Repp, 1982). Following Catford (1977), this coordination pattern is called close transition, which is defined as reaching the target of the second consonant in a cluster before the constriction of the first consonant is released. Close coordination between two consonants would entail that the release of the first

consonant is overlapped by the target of the second consonant. This coordination relationship is illustrated in Fig. 4.

----------------------------Insert Figure 4 about here ------------------------------

Another type of consonant coordination is demonstrated by Moroccan Colloquial Arabic (MCA). Gafos hypothesizes that the transitional schwa in coda clusters (e.g. [smim^{ə}n]) arises from a timing relationship in which there is not a continuation from the target plateau of the first consonant to that of the second. This leaves a short period of open vocal tract between the two consonants, which gives rise to the transitional schwa. The type of consonant coordination found in MCA is demonstrated in Fig. 5.

----------------------------Insert Figure 5 about here ------------------------------

In order to analyze different coordination patterns found in MCA, Gafos (2002) develops an Optimality Theoretic grammar that includes constraints on how consonant gestures may be aligned with respect to one another. However, Gafos only discusses consonant clusters that are allowable in MCA, and therefore his analysis does not include constraints that distinguish between consonant sequences that are attested in MCA and those that are not. In general, this is true for most formal accounts of coordination (e.g. Bradley, 2002; Cho, 1998), except Hall (2003), who proposes some phonotactic constraints to account for transitional vowels in obstruent-sonorant clusters. In the analysis of the English production facts, a set of gesturally-based phonotactic constraints are proposed to account for the English speakers' performance on obstruent-initial clusters. Whereas the phonotactic constraints proposed by Hall (2003) employ traditional labels like "obstruent" and "liquid", the current analysis aims to develop a cohesive

phonological grammar that makes reference to both the spatial gestural primitives and temporal relationships of AP.

The analyses of languages such as MCA present an alternative to the assumption that the schwa produced by an English speaker in a word like [z$^{\circ}$gano] is the epenthesis of a phonological vowel. A gestural coordination account suggests that manipulation of the alignment of the gestures is also available in the production of non-native sequences. Thus, an English speaker presented with a word like *ftalu* or *zgano* who repairs the initial consonant sequences by inserting some kind of schwa-like material may be exhibiting one of two repair possibilities, as demonstrated by the gestural schematics in Fig. 6. The case where the consonants are not overlapped is referred to as "gestural mistiming".

------------------------------Insert Figure 6 about here ------------------------------

The question of how speakers deal with non-native sequences in production is an important one, since it helps to establish the interrelation between the linguistic elements that speakers require for speech production. That is, Articulatory Phonology is partially responsible for counteracting the assumption implicit in phonological theory that the articulatory coordination of sequential sounds is not within the domain of phonology, but rather at a motoric level that implements the phonological plan. When speakers produce non-native sequences with various rates of success and with different coordination than they use for sequences attested in their native language, it demonstrates that phonotactics—a classic phonological phenomenon—interacts with articulatory coordination and provides evidence that there is not necessarily a neat division between phonological planning and articulatory implementation.

3. Phonetic characteristics of the production of non-native sequences

The epenthesis and gestural mistiming repair possibilities introduced in Fig. 6 make two different predictions with regard to acoustic output. If speakers are epenthesizing a vowel, then it should have the same acoustic characteristics as a schwa that appears lexically in the same environment. On the other hand, if speakers are mistiming gestures and not producing a schwa, then a speaker's vocal tract during the open period will not have the same configuration as it would for a vowel. Articulatorily, a sequence repaired by mistiming the gestures should appear more similar to a consonant cluster than to a /CəC/ sequence. Two studies examining the acoustics and articulation of non-native sequences are reviewed in this section.

3.1. Acoustic differences between lexical and inserted schwa

The first study compares the acoustic properties of a lexical schwa with the transitional schwa English speakers inserted when attempting to produce non-native fricative-initial clusters (FC) (Davidson, 2006). The stimuli were composed of the initial consonants /f, z, v/ in combination with obstruents (/p, t, k, s/ for /f/-initial words and /b, d, g, v, z/ for /z/- and /v/-initial words) and nasals (/n, m/) to make corresponding pseudoword stimuli of the form CCaCV and CəCaCV. For example, speakers were presented with stimuli such as /fmasa/ ~ /fəmasa/, /zgano/ ~ /zəgano/, /vzagi/ ~ /vəzagi/. There were 24 CC stimuli and 24 CəC stimuli for each initial fricative. The stimuli were presented in random order both orthographically on a computer screen in English-like orthography (e.g. "vzagy") and aurally, pronounced by a Slovak speaker. Twenty participants repeated each word aloud once. Using a spectrogram, tokens were coded either as correct, or for the repair exhibited. In addition to tracking repair types, the duration and

first and second formant frequencies of both the inserted and lexical schwas were measured (all experimental details are extensively presented in Davidson, 2006).

Results for the sequences are shown in Fig. 7 and the coding scheme is shown in Table 1. Two aspects of these findings are important for this discussion. First, the proportion of correctly produced tokens demonstrates that speakers are not equally accurate on all non-native initial sequences. Speakers are significantly more accurate on /f/-initial words (i.e. /fC/), followed by /zC/ and then /vC/. The second important aspect of the results is that for all sequence types, speakers are most likely to repair the non-native cluster with insertion.

-----------------------------Insert Figure 7 about here ------------------------------

-----------------------------Insert Table 1 about here ------------------------------

While other repairs are also attested, the insertion repair is particularly interesting, since it is the most common and its underlying origin is ambiguous. To determine whether the inserted material is due to epenthesis or gestural mistiming, acoustic characteristics of both the inserted material and the lexical vowel in corresponding sequences were examined. These results are shown in Table 2, which contains the duration, first formant (F1) and second formant (F2) midpoint values of both inserted and lexical schwas for each type of sequence. All the differences between lexical and inserted schwas are statistically significant, except for the F2 difference for /fC/. The findings indicate that the inserted schwas do not share the same acoustic properties as the lexical schwas, which is consistent with gestural mistiming. Specifically, a lower F1 for all sequence types indicates that the inserted schwa is being produced with a greater constriction than the lexical schwa, and a lower F2 indicates greater coarticulation with, or

11

anticipation of, the following /a/ than a lexical vowel would have (each non-native cluster was followed by /a/ in the stimuli).

------------------------------Insert Table 2 about here ------------------------------

3.2. Articulatory differences between lexical and inserted schwa

In addition to the acoustic findings, an ultrasound study comparing the production of lexical and inserted schwas provides converging articulatory evidence for gestural mistiming (Davidson, 2005). To further investigate the epenthesis and gestural mistiming hypotheses, the production of non-native /zC/ sequences was compared to the production of legal clusters beginning with /s/. /sC/ and /zC/ sequences are equivalent in their oral articulation, with the difference between them being the voicing for /z/. First, the production of /zC/ sequences was compared to the production of /sC/ clusters that were matched for the place, manner, and continuancy of the second segment. Second, the production of /zC/ sequences was also compared to the production of /səC/ sequences which contain an underlying lexical schwa.

The ultrasound images were used to compare tongue shapes when speakers' produced /zC/ as [zᵊC] to their production of /sC/ clusters and /səC/ sequences. It was hypothesized that if tongue shape changes during the production of the sequence /zC/ are more like /səC/ than /sC/, it would suggest that there is phonological epenthesis of a schwa gesture into the /zC/-initial word. If, on the other hand, tongue shapes for /zC/ are more similar to those for /sC/, despite the acoustic presence of a schwa, it would suggest that the schwa percept arises from gestural mistiming. The ultrasound images of five native speakers of American English producing three triads of /sC-/, /səC-/, and /zC-/ initial words were recorded (e.g. superfluous ~ spurt ~ *zbura*; satirical ~ steer ~ *zdiri*; succumb ~ scum ~ *zgama*). Seven repetitions of each word were

recorded, the middle five of which were used in the statistical analyses. Ultrasound has a resolution of 33 frames per second, so the first five frames corresponding to the start of the fricative and going to the closure of the stop were compared.

Results from the ultrasound study showed that the [zᵊC] productions of three of the five speakers were articulatorily more similar to [sC] than to [səC]. That is, for these speakers, although targets like /zgama/ were produced as [zᵊgama], there was no evidence of a phonemic schwa in the tongue shapes as the tongue moved from [z] to the following consonant. A root mean squared error metric indicated that the ultrasound frames corresponding to the production of the [zᵊC] sequence were consistently more similar to the frames for the [sC] sequence.

The other two speakers did not quite show this pattern. One speaker's tokens were more consistently like /səC/, and the other speaker's productions were intermediate between /səC/ and /sC/. This is likely because her tokens were more variable, and more data would have been necessary to discern her pattern. The fact that speakers did not all behave the same is not surprising, as the second language phonology literature often reports individual differences in how speakers produce non-native sequences (e.g. Abrahamsson, 2003; Broselow & Finer, 1991; Hansen, 2004). The remainder of the analysis focuses on a theoretical account of the majority of the speakers, who repaired non-native sequences with gestural mistiming.

In the next section, a phonological account of the speakers' behavior on non-native sequences is developed. Following Gafos (2002), coordination relationships among adjacent segments are taken to be governed by phonological constraints. Furthermore, phonotactic restrictions are implemented through a ban on certain types of coordination between particular gestures. The full grammatical analysis developed in Section 4 accounts for speakers' behavior on those sequences usually considered non-native.

4. Temporal coordination, phonotactics, and variation in the production of initial sequences

In this section, an Optimality Theoretic grammar integrating phonotactics and coordination is developed. Optimality Theory (OT) is an ideal tool for building a representationally coherent framework that combines phonotactics with the gestural primitives of Articulatory Phonology; the theoretical mechanisms of OT are agnostic about the representational elements best suited to phonological systems, but the fundamental insight of OT—constraint ranking—can account for the relationship between phonotactics and coordination both in second language acquisition and in the world's languages. The tenets of OT are laid out in Smolensky's (2006) article in this volume, and also in a previous paper in *Cognitive Science* (Smolensky, 1999:596):

   a. Given an input, the grammar produces as output the linguistic structure that maximizes Harmony.
   b. The Harmony of a potential output is the degree to which it simultaneously satisfies a set of violable constraints on linguistic well-formedness (including constraints requiring that the output faithfully express the input).
   c. The constraints have different strengths, determining which take priority when constraints conflict.
   d. The grammar of a language is a ranking of constraints from strongest to weakest; a higher-ranked constraint has absolute priority over all lower-ranked constraints.
   e. The set of possible outputs, and the set of constraints, is the same in all languages; grammars of languages differ only in the way constraints are ranked.

OT posits that grammar is composed of a set of constraints on the output that evaluate the possibilities for going from an underlying representation to a surface form. One advantage of OT is that grammars change by reranking constraints, which can account for both typological differences between languages and for differences in production among speakers of the same language. The ranking of constraints further provides a means for understanding the process of

14

language acquisition, and how learners' native languages affect the production of either a second language or non-native sequences.

4.1. Coordination in a gestural grammar

The idea that coordination between adjacent gestures is determined by the phonology was originally proposed by Browman and Goldstein (1986 et seq.), and formalized in an OT grammar by Gafos (2002) for Moroccan Colloquial Arabic. In OT, coordination is governed by a set of ALIGNMENT constraints (McCarthy & Prince, 1993). The generalized constraint is given in (2).

(2)   ALIGN(G_1, landmark$_1$, G_2, landmark$_2$): Align landmark$_1$ of gesture G_1 to landmark$_2$ of gesture G_2

For consonants in close transition (as illustrated in Fig. 4 above), an alignment constraint ensuring that the release of the first consonant is coordinated with the target of the second consonant is posited. While English speakers initially do not produce the Slavic fricative-initial targets with close transition, this is ultimately the target coordination pattern if they were to produce these sequences as Slavic speakers do. Furthermore, this is the same coordination pattern that speakers use for /sC/ sequences that are attested in English. The constraint for this coordination relationship is given in (3); the landmarks refer to those exemplified in Fig. 3.

(3)   ALIGN (C_1, release, C_2, target): Align the release of gesture C_1 to the target of gesture C_2

ALIGN constraints govern the type of coordination allowed between two consonants, but they do not mediate the phonotactics of a given language. That is, the ALIGN constraints must interact with a set of constraints that determines which consonants in a language can be sequentially ordered. In the next section, evidence is first presented to further confirm the

phonological origin of the English experimental results. Second, a demonstration of a constraint hierarchy containing both ALIGN and simplified phonotactic constraints shows how languages can differ as to their consonant cluster inventories, and how gestural mistiming might arise on the surface when speakers attempt to produce non-native sequences. The actual form of the phonotactic constraints is then presented in Section 4.3.

4.2. The cross-linguistic status of FC sequences

The results from the experiment discussed in Section 3.1 showed that English speakers producing non-native sequences beginning with /fC/, /zC/, and /vC/ are not equally accurate on all of the stimuli, despite the fact that none of them are permissible in English. Specifically, speakers are most accurate on /fC/ sequences (65%), followed by /zC/ (44%) and then /vC/ sequences (27%). Davidson (2006) shows that these results are not correlated with the frequency of such sequences in word-medial or word-final position. Furthermore, they cannot be attributed to articulatory factors alone. For example, many of the experimental sequences are possible word-medially and finally in English (e.g., Ma[zd]a, hu[zb]and, lo[vb]ird, ra[ft]), but speakers are not able to transfer this articulatory knowledge to word-initial position. In other words, one hypothesis regarding the English speakers' performance is that their poor production of unfamiliar sequences is due to motoric inexperience. If this were the case, it would be expected that speakers would be more accurate on those sequences which are attested in other parts of the words as compared to those sequences that are not. However, not only do speakers have trouble transferring the motoric patterns that they utilize for word-medial and final sequences to initial position, but their production of sequences which are attested (e.g. tran[zg]ression) are no more

16

accurate than sequences which are not attested in English in any part of the word (e.g. [vg]) (Davidson, 2006).

If the English speakers' performance cannot be attributed to motoric inexperience, then its origin must lie elsewhere. The phonological grammar is a good candidate. Already English contains phonotactic restrictions on /fC/, /zC/, and /vC/ sequences existing in word-initial position since there are no such English lexical items. Furthermore, there is reason to believe that each of these sequences has a separate phonotactic constraint that can be ranked in the grammar. The cross-linguistic distribution of these FC clusters reveals that there are languages that have only /fC/ clusters, only /zC/ clusters, or all of /fC/, /zC/, and /vC/. For example, Norwegian and Afrikaans have /fC/ sequences (e.g. Afrikaans *fnuik* 'clip the wings, frustrate'), Hebrew, Croatian, and Romanian have only /zC/ clusters (e.g. Croatian *zveka* 'sound'), and Greek, Tsou, and several Slavic languages have all three types of fricative clusters (e.g. Polish *[ft]orek* 'Tuesday', *[zb]adać* 'explore', *[vz]ad* 'back'). In order to define inventories that have one but not all of these sequences, there must be individual constraints prohibiting the other consonant clusters. Following the basic tenets of OT, it is hypothesized that these constraints are universal, and therefore can be manipulated by English speakers producing non-native sequences.

The basic OT machinery necessary to capture the difference between the inventories of languages that allow FC clusters and those that do not is illustrated in (4). The ALIGN constraint for close transition, which pertains to FC sequences in both Polish and English, for example, is ranked differently with respect to phonotactic constraints in each language. In Polish, where all of the experimental initial fricative clusters are allowed and are produced with close transition, ALIGN is ranked above constraints that would prohibit /fC/, /zC/, and /vC/ from having close

17

coordination. For the time being, the phonotactic constraints are indicated with the placeholders *fC, *zC, and *vC; the actual gesturally-based constraints are developed in Section 4.3.

Given the input /zgano/ first introduced in Fig. 2, there are many ways that the surface form could be produced. In (4), just two possibilities are temporarily considered: [zgano], the candidate that is faithful to the input, and [z°gano], the candidate in which the target of C_2 does not overlap the release of C_1, leading to an open transition between the consonant. In OT, the violation of a higher ranked constraint, specified by the marker "*!" in the tableau in (4), is called a *fatal violation*. The violation of such a high-ranked constraint indicates that the corresponding surface form (here (4)b), cannot be the actual one. Instead, even though candidate (4)a violates *zC, it is the correct output for Polish. The winning candidate is indicated with the symbol '☞'. Because /fC/, /zC/ and /vC/ are all allowed in Polish, these constraints are all ranked below the ALIGN constraint and there is no evidence that they are ranked with respect to one another. The lack of ranking among the phonotactic constraints is shown by the dotted lines in the tableau.

(4) Polish: /zgano/	ALIGN (C_1, release, C_2, target)	*vC	*zC	*fC
a. ☞ [zgano]			*	
b. [z°gano]	*!			

In English, the opposite of Polish is true: all of the phonotactic constraints are ranked above ALIGN. As the experimental results reviewed in Section 3 demonstrated, English speakers faced with /fC/, /zC/ and /vC/ initial sequences repair them by producing the consonants with a non-overlapping coordination. If the phonotactic constraints are formulated so as to prohibit

close coordination between the consonants, then another way to repair a sequence like /zg/ would be to insert a vowel between the consonants. However, since this is not what English speakers do, a constraint banning the insertion of a gesture into a surface form must be ranked above the ALIGN constraint. This constraint, called DEP (McCarthy & Prince, 1999), is defined in (5).

(5)　DEP: Every segment in the output has a correspondent in the input. (Do not insert a gesture.)

The tableau for English, including the new constraint DEP, is shown in (6). Because violations of *zC and DEP are both fatal, this data is not sufficient for determining whether DEP is crucially ranked with respect to the phonotactic constraints.

(6)　English: /zgano/	DEP	*vC	*zC	*fC	ALIGN (C_1, release, C_2, target)
a. [zgano]			*!		
b. ☞ [zᵊgano]					*
c. [zəgano]	*!				

In the next section, the placeholders *fC, *zC and *vC are replaced with phonetically motivated, gesturally-based constraints. These constraints set the stage for explaining how English speakers manipulate their phonological systems in order produce variable performance on non-native FC sequences, which will be discussed in Section 4.4.

4.3. Phonotactics in a gestural grammar

Cross-linguistically, word-initial obstruent clusters are uncommon. In English, non-strident fricatives are disallowed in initial clusters when the second member of the cluster is a

non-approximant (e.g. */f v θ ð/+stop, fricative, nasal) (Hammond, 1999). This observation

follows from the claim that obstruents prefer to release into vowel-like approximants, such as

liquids or glides, since the resultant formant transitions provide more cues as to the place of

articulation of the fricative (Côté, 2000; Kingston, 1985; Steriade, 1993, 1997). In order to

improve the perceptual environment for non-strident FC clusters, the minimally restrictive

solution is to ensure that the release of the first consonant is not obscured by the plateau of the

second consonant. Although it may be best to be followed by an approximant (including vowels),

at the very least the release of an obstruent provides further cues to the identity of the consonant

(e.g. Blumstein & Stevens, 1978; Dorman, Studdert-Kennedy, & Raphael, 1977; Kingston,

1990). Thus, initial /f+obstruent/ or /f+nasal/ may be ruled out by a prohibition on the overlap of

the release of the /f/ gesture by the plateau of the following obstruent or nasal gesture.

To ensure that the release of one consonant is not obscured by the following consonant, a

family of constraints is proposed to ban this type of overlap for two segments that are specified

for particular gestural parameters. The general version of this constraint, called RELEASE/Gα,Gβ

is defined in (7):[1]

(7)   RELEASE/Gα,Gβ: Do not overlap the release of a consonant containing gestural
        parameter(s) Gα with the plateau of a following consonant containing gestural
        parameters(s) Gβ.

---

[1] The RELEASE constraint as defined in (7) must actually be part of a family of constraints that distinguishes between

different positions in the word. Many languages, including English, allow clusters like the experimental ones in

word-medial position while they do not allow them in initial position (e.g. *zbagi, but fri[zb]ee). It has been argued

that there are phonetic reasons for such positional differences (e.g., Steriade, 1997). As currently worded, the

RELEASE constraints do not differentiate position, but fully fleshing out this constraint family is beyond the scope of

the paper. For the purposes of this discussion, it should be kept in mind that these constraints are meant to pertain to

word-initial sequences, and not necessarily to medial ones.

In a phonological system that employed traditional binary features, it would be easy to ban /fC/ sequences: *[-strident][-approximant] (cf. Morelli, 1999). [-strident] refers to the class of low-intensity (non-strident) fricatives such as /f v θ ð/, and [-approximant] pertains to stops, fricatives, and nasals. This prohibition is more difficult to describe in Articulatory Phonology, because there are currently not gestural parameters that differentiate between the classes of high-intensity (strident) fricatives /s z ʃ ʒ/ and low-intensity fricatives /f v θ ð/ in English. Consequently, new elements that are compatible with Articulatory Phonology's gestural parameter values are proposed in (8) in order to capture these groupings. [CD=CRITICAL NO-GROOVE], the Gα constituent, specifies non-strident fricatives. [Critical] is the AP parameter for the constriction degree (CD) corresponding to fricatives, and [No-groove] refers to fricatives that are not produced with a grooved tongue. Compared to grooved strident sounds like /s/ or /ʃ/, these fricatives are much less perceptible, especially at word edges (Kingston, 1990). [CD<Narrow], the Gβ constituent, is a way to group obstruents (including nasal stops), or gestures with the constriction degrees [Closure] (stops) and [Critical] (fricatives). Other constriction degree parameters in Articulatory Phonology—Narrow, Mid, and Wide—correspond to approximants and vowels.

(8)   [CRITICAL NO-GROOVE] = Gestures with a constriction degree of [Critical] that are not produced with a grooved tongue.
[CD<NARROW] = Gestures with a constriction degree that is more constricted than [Narrow].

For a sequence like /ft/ (or any FC cluster beginning with [f v θ ð]), the RELEASE constraint can be formulated as in (9).

(9)   RELEASE/[CRITICAL NO-GROOVE],[CD<NARROW]: Do not overlap the release of a non-strident fricative with the plateau of a non-approximant. (For the sake of readability, this will be abbreviated as REL/[NoGroove],[NonApprox])

21

Whereas the production of a transitional schwa between /f/-initial sequences provides a better perceptual environment for consonants with weak cues to their identity, /zC/ sequences benefit from a transitional schwa for aerodynamic reasons. The optimal situation for obstruent voicing occurs when oral pressure is maximally lower than glottal pressure. However, for fricative clusters, frication is facilitated when oral pressure is maximally higher than atmospheric pressure (Ohala, 1994). This sets up conflicting articulatory requirements for the production of voiced fricatives. Voiced obstruent clusters are especially disadvantaged by the fact that they are longer in duration than single voiced obstruents, which may require the conflicting air pressure requirements to be sustained for longer than the speech system can accommodate (Ohala & Kawasaki-Fukumori, 1997; Westbury & Keating, 1986).

In Articulatory Phonology, it has been proposed that the main "articulator" concerning voicing in obstruents is the volume of the pharyngeal cavity, since it is the manipulation of the volume in this region which ensures the appropriate transglottal pressure for the maintenance of voicing during obstruents (McGowan & Saltzman, 1995). In fact, the expansion of the pharyngeal cavity necessary for the production of voiced obstruents is the result of a number of factors, including expansion of the pharyngeal walls and lowering of the larynx (Bell-Berti, 1975; Ohala, 1994; Westbury, 1983). Since these actions all serve the goal of increasing the volume of the supraglottal cavity, McGowan and Saltzman consolidate them into a single aerodynamic tract variable for the purpose of simulations.

Since the primary role of the supraglottal cavity volume is to maintain voicing on obstruents, the gestural parameter most relevant to a RELEASE constraint prohibiting voiced

obstruent clusters can be labeled as [PHARYNGEAL VOLUME]. The constraint regarding /z/-initial

clusters is shown in (10).

(10) RELEASE/[PHARYNGEAL VOLUME],[CD<NARROW]: Do not overlap the release of a voiced
obstruent with the plateau of a non-approximant. (REL/[VoicedObs],[NonApprox])

Given the definitions in (9) and (10), /v/-initial clusters cause violations of both of these

constraints since the clusters are both non-strident and voiced. However, the experimental results

demonstrate that the two constraints already proposed are not sufficient to account for

performance on /vC/ sequences, since speakers are significantly less accurate on /vC/ clusters

than they are on either /fC/ or /zC/. This performance suggests that speakers make a three-way

distinction between the different cluster types, and subsequently, that three different constraints

are necessary to account for the findings. In fact, it is not surprising that speakers perform worse

on /vC/ sequences since /v/ is both a weak-intensity fricative and also voiced. /vC/'s violation of

both the RELEASE/[NoGroove],[NonApprox] and RELEASE/[VoicedObs],[NonApprox]

constraints is a worst-of-the-worst situation that distinguishes it from the other fricatives.

Inventories that ban only the worst-of-the-worst are common and are dealt with in OT with

locally conjoined constraints (Kirchner, 1996; Lubowicz, 1998; Moreton & Smolensky, 2002;

Smolensky, 2005). A local conjunction is a constraint formed through the combination of two

lower-ranked constraints that have in common the same domain of application. When each of

these constraints is violated separately, their violations are not enough to be fatal, but when the

conjoined constraint is violated, the candidate which violates it cannot be optimal. /vC/ clusters

can be ruled out by conjoining the two lower-ranked RELEASE constraints. The conjoined

constraint is shown in (11).

(11) RELEASE/[CRITICAL NO-GROOVE],[CD<NARROW]&RELEASE/[PHARYNGEAL VOLUME],
[CD<NARROW]: Do not overlap the release of a non-strident voiced fricative with the

plateau of a non-approximant.
(REL/[NoGroove],[NonApprox]&REL/[VoicedObs],[NonApprox])

Now that all of the phonotactic constraints have been defined, the grammar that is
assumed for a monolingual English speaker is given in (12) (here without DEP). The notation *(!)
indicates that for /vC/ sequences, any of the violations of the constraints could be the fatal one
since all of the violated constraints are ranked above ALIGN.

(12) /fkada/	REL/[NoGroove], [NonApprox]& REL/[VcdObs], [NonApprox]	REL/[Vcd Obs],[Non Approx]	REL/[No Groove], [NonApprox]	ALIGN(C_1, release, C_2, target)
a. fkada			*!	
b. ☞ fᵊkada				*
/zgano/				
c. zgano		*!		
d. ☞ zᵊgano				*
/vzagi/				
e. vzagi	*(!)	*(!)	*(!)	
f.☞ vᵊzagi				*

In addition to demonstrating how FC clusters are prohibited in the cluster inventory of
English, the constraints in (9)-(11) play a crucial role in accounting for the variability exhibited
in the results in experimental production. A preview of the ranking of RELEASE constraints
necessary for capturing the production facts is already applied in (12), and is shown again in
(13). A more detailed exposition of these rankings and how they reflect the experimental
production facts will be given in the next section.

(13) REL/[NoGroove],[NonApprox]&REL/[VoicedObs],[NonApprox] ≫
    REL/[VoicedObs],[NonApprox] ≫ REL/[NoGroove],[NonApprox]

24

4.4. Variation in OT grammars: accounting for non-native production

A number of studies have demonstrated that variation can arise in phonological processes (Anttila, 1997; Boersma, 1998; Boersma & Hayes, 2001; Davidson, Jusczyk, & Smolensky, 2004; Nagy & Reynolds, 1997; Reynolds, 1994; Zuraw, 2003). In this situation, a constraint may be highly ranked some proportion of the time, giving rise to one phonological output, while the remainder of the time the constraint is lower-ranked and another output is attested. Anttila (1997), for example, accounts for optional use of different allomorphs for the Finnish genitive plural in a single phonological environment. Several formal analyses of phonological and syntactic variation have been presented within a version of Optimality Theory which employs *floating constraints* (Anttila, 1997; Davidson & Goldrick, 2003; Davidson et al., 2004; Davidson & Legendre, 2003; Nagy & Reynolds, 1997; Reynolds, 1994). A floating constraint situation occurs when certain constraints do not have a fixed rank with respect to other strictly ranked constraints, giving rise to multiple outputs. In this section, a ranking containing a floating constraint is developed to account for the non-native cluster production experiment results.
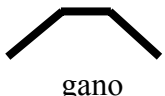
In the production experiment, speakers attempted to accurately produce the target clusters containing /fC/, /zC/, and /vC/ clusters. When they failed to do so, they most often repaired the sequences with gestural mistiming. Ultimately, ranking the appropriate alignment constraint higher than all RELEASE constraints is necessary if a speaker is to produce each of these clusters accurately on each attempt. However, as demonstrated in the OT child language acquisition literature, learners do not usually change a ranking so dramatically with little data, so one way to accomplish the goal is to posit a floating range for a constraint (e.g. Davidson & Goldrick, 2003; Legendre, Hagstrom, Vainikka, & Todorova, 2001). If the speakers in the experiment were actually learning a Slavic language that contained the target clusters, they may allow ALIGN to

float over the RELEASE constraints as part of a gradual learning mechanism, much like infant learners appear to do.

Speakers who repair unattested clusters with gestural mistiming have already ranked DEP, which was introduced in (5), and the constraints in (14) above the RELEASE constraints.
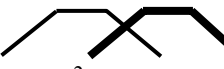
(14) a. MAX: Every segment in the input has a correspondent in output. (Do not delete a gesture.)
b. IDENT: A segment containing a feature [γF] in the input should correspond to in the [γF] output. (Do not modify the constriction degree or constriction location of a gesture). (McCarthy & Prince, 1999)

The tableau in (15) is a more complete version of the ranking for an English speaker who repairs FC clusters with gestural mistiming. The candidates (15)c, d, e violate DEP, MAX, and IDENT, respectively, which are all higher ranked than both REL/[VoicedObs],[NonApprox] and ALIGN(C_1, release, C_2, target). There is no discernable ranking between these constraints.

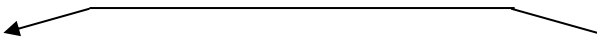| (15) /zgano/	DEP	MAX	IDENT	REL/[VoicedObs], [NonApprox]	ALIGN(C_1, release, C_2, target)
a. zgano				*!	
b. ☞ zᵊgano					*
c. zəgano	*!				
d. gano		*!			
e. sgano			*!		

The existence of other possible repairs as illustrated in (15) raises the question of why many experimental participants seem to employ gestural mistiming as the preferred mechanism for producing the unattested clusters. One explanation for the commonality of this repair is that it maximizes the recoverability of the intended sequence. Researchers have suggested that when non-native speakers are aware of the intended phonemes of the word they are attempting to produce, they prefer to preserve as much information as possible (e.g., Abrahamsson, 2003; Weinberger, 1994). This preference is violated by deleting or changing a gesture, explaining why MAX and IDENT are ranked above the RELEASE constraints. Though epenthesis also preserves the intended consonants, the high ranking of DEP may be due to a dispreference for placing additional structure (e.g. a new gesture) into an underlying form. As noted earlier, learners acquiring a new language must learn not only which sequences can form a unit in the target language, but also what the appropriate coordination for those sequences is. During the experiment, the participants may realize that like English, the stimuli also have close coordination, and so they will be accurately produced when ALIGN(C_1, release, C_2, target) is ranked above all of the RELEASE constraints.

Speakers attempting to faithfully reproduce an /z/-initial cluster under experimental conditions in fact sometimes successfully rerank ALIGN above the phonotactic RELEASE constraints prohibiting the experimental cluster in English. This is shown in (16).

(16) /fkada/	ALIGN(C_1, release, C_2, target)	REL/[VoicedObs], [NonApprox]	*ALIGN(C_1, release, C_2, target)*
a. ☞ zgano		*	
b. zᵊgano	*!		

The experimental results indicate that REL/[NoGroove],[NonApprox] is the lowest ranked of the RELEASE constraints, followed by REL/[VoicedObs],[NonApprox], with REL/[NoGroove],[NonApprox]&REL/[VoicedObs],[NonApprox] ranked above both simple constraints. The ranking of the constraints is determined by speakers' accuracy on each type of cluster. The previous tableau shows that if ALIGN is ranked above REL/[NoGroove], [NonApprox], then /fC/ sequences will be accurately produced. But if instead the ALIGN constraint is ranked over both REL/[NoGroove],[NonApprox] and REL/[VoicedObs], [NonApprox], a speaker will accurately produce both /fC/ and /zC/ words, but not /vC/. This is shown in the tableau in (12) (gestural schematics are omitted for conciseness). By allowing the ALIGN constraint to float over the whole range of RELEASE constraints, the gradient accuracy of speakers on /fC/ > /zC/ > /vC/ sequences can be explained.[2]

---

[2] It is because the ALIGN constraint appears to float over the entire range of RELEASE constraints that floating constraints are the appropriate theory to account for the performance in the Davidson (2006) experiment. An alternative theory, stochastic Optimality Theory, has also been proposed to account for variation (Boersma, 1998; Zuraw, 2000), but it is not suited to the experimental data examined in this paper. In stochastic OT, constraints have distributions, and in cases where the distributions of two constraints overlap, they can sometimes be reranked in the grammar. However, the range over which constraints can overlap is restricted by the stipulation that all constraints have the same probability distribution. This means that it is impossible in stochastic OT for one constraint to have a

(17) /fkada/	REL/[NoGroove], [NonApprox]& REL/[VcdObs], [NonApprox]	ALIGN(C_1, release, C_2, target)	REL/[Vcd Obs],[Non Approx]	REL/[No Groove],[Non Approx]	*ALIGN(C_1, release, C_2, target)*
a.☞ fkada				*	
b. fᵊkada		*!			
/zgano/					
c.☞ zgano			*		
d. zᵊgano		*!			
/vzagi/					
e. vzagi	*!		*	*	
f.☞ vᵊzagi		*			

Without data from a task in which speakers are faced with phonotactically unattested forms, the "hidden" ranking of the RELEASE constraints would not be evident. The ranking is referred to as hidden since no lexical items of English contain these FC sequences word-initially in English and there would not otherwise be evidence that English speakers distinguish between /fC/, /zC/, and /vC/ word-initial sequences. Speakers who correctly produce non-native clusters at least some proportion of the times that they attempt them are able to move the coordination constraint out of the position it occupies in the native English grammar, which ultimately uncovers the hidden rankings.

The next question this leads to, then, is what the idea of "percent correct" on the experimental targets corresponds to in phonological terms. One possibility is that at the

distribution so large that it can overlap with the distributions of three constraints that are all ranked with respect to one another. While stochastic OT might be appropriate for the variation seen in the final state of the grammar, it does not appear to accurately capture learning situations (see also the data in Davidson & Goldrick, 2003; Davidson & Legendre, 2003).

beginning of the task, speakers rerank ALIGN to some position in the hierarchy and allow it to remain there for the duration of the experiment. However, this predicts that they should be performing at 100% for those clusters whose corresponding RELEASE constraints are ranked below the coordination constraint, while never accurately producing those clusters governed by higher ranked constraints. In fact, speakers do not show all-or-none performance, suggesting that they are not reranking the coordination constraint at the beginning of the experiment and leaving it in that position. Instead, speakers are more likely reranking the coordination constraint spontaneously for each trial. Such reranking is possible because ALIGN is allowed to float over the range of RELEASE constraints. For each optimization, ALIGN is assigned a fixed position in the hierarchy. This position is potentially different for each attempt of a target cluster. Some clusters, but not others, will be possible depending on where the coordination constraint lands each time it is reranked. The floating range for ALIGN is demonstrated in Fig. 8.

----------------------------Insert Figure 8 about here -------------------------------

The fact that ALIGN can float to any position in the hierarchy predicts that if it floats above a higher constraint some proportion of the time, clusters banned by a lower constraint must be accurately produced even more often. The position marked by ❶ defines the base grammar of English, which disallows all of the experimental sequences. By reranking ALIGN to position ②, speakers correctly produce /fC/ clusters, whereas elevating ALIGN to position ③ adds to the inventory /zC/ clusters. A final elevation of ALIGN leads to the inclusion of /vC/ clusters. The simplest hypothesis regarding the floating range of ALIGN is that of "uniform floating", in which each of the positions ❶-④ is an equally likely resting place (Anttila, 1997). Uniform floating entails that clusters that can be attained by a greater number of docks should be

accurately produced more often. For example, if there is equal probability that the alignment constraint will land in any of the positions whenever an illegal cluster is attempted, the /fC/ clusters should be correctly pronounced approximately 75% of the time, since three of the four positions allow them.

Uniform floating does not seem to hold up when the patterns of individual speakers are examined. However, though speakers do not necessarily conform to the proportion correct predicted by the uniform floating hypothesis for each cluster stratum ❶-④, the performance of all 20 speakers is nevertheless consistent with the fixed ranking of constraints given in (13). In other words, assuming that all English speakers share this common hidden ranking of constraints, performance on /fC/ clusters statistically equals or exceeds performance on /zC/ clusters, which equals or exceeds performance on /vC/ clusters for all of the participants. In Table 3, the percentages predicted by uniform floating are shown next to the observed experimental percentages for one participant, Speaker 8, who preferred insertion for all but 1% of his repairs. Determining the floating ranges for ALIGN for each individual speaker is not the goal of this analysis, but it should be noted that differences among speakers may be attributable to speakers' varying ability to allocate the cognitive resources necessary to elevate the coordination constraint (see Davidson et al., 2004). Furthermore, non-uniform floating is not limited to experimental production; it can also be found in child language acquisition (Davidson & Goldrick, 2003).

-----------------------------Insert Table 3 about here ------------------------------

4.5. The origins of hidden rankings

A hidden ranking analysis of any data, whether experimental or in language acquisition, inevitably raises the question of the origin of these hidden ranking. There are two possibilities regarding the basis for hidden rankings. The first possibility is that the ranking is present in the initial state. In this case, the constraint hierarchy in (13) can be considered a default ranking that can change given the appropriate input during first language acquisition. Children acquiring languages like Hebrew or Serbo-Croatian then will learn that the ranking

REL/[VoicedObs],[NonApprox] ≫ REL/[NoGroove],[NonApprox] must be reversed. This possibility also predicts that speakers of languages that do not have the relevant initial clusters should show the same pattern as English speakers on the experiment in Section 3.1, since they would have had no impetus for changing the default ranking. This could be easily tested by administering the experiment to speakers of languages like Spanish, Chinese, or Hindi.

Another possibility is that such rankings are language-specific. One way language-specific rankings might occur is as a result of the learning process (e.g. Boersma & Levelt, 2000; Tesar & Smolensky, 2000), which could end up leading to particular hidden rankings, even if they will never have an effect on the base English phonology. As discussed in Davidson (2006), another factor that may affect language-specific hidden rankings is analogical generalization from existing sequences. (see Baayen, 2003; Bybee, 2001; Skousen, 1989 for similar ideas regarding generalization across morphologically related words). That is, /fC/ sequences are preferred because some /f/-initial clusters (such as *flight* or *fright*) are already allowed in English. While there are no allowable /z/-initial sequences, speakers may draw on their knowledge of /sC/ —another coronal fricative-initial sequence that crucially can combine with nasals and obstruents—to produce these targets. This link to a legal sequence is somewhat less direct than that for /f/-initial sequences, but /vC/ sequences are even further removed from existing clusters

in that /v+obstruent, nasal/ does not get the same direct benefit from /f/ as /z+obstruent, nasal/ does from /s/, since there are no attested /f+obstruent, nasal/ sequences. If analogy is a mechanism that can affect the shape of the grammar as researchers like Baayen (2003), Bybee (2001), and Skousen (1989) have posited, then the ranking of the constraints may reflect this. A complete account of the relationship between analogy and Optimality Theoretic grammars is left to future research.

5. Conclusion

As a step toward integrating insights from Articulatory Phonology and previous Optimality Theoretic analyses of phonotactics, the perceptual, articulatory and aerodynamic factors that determine preferred phonotactic structures and the temporal coordination of gestures must be incorporated into phonological theory. A framework that takes the gestures of Articulatory Phonology to be the basic units of representation must be capable of both representing phonotactic restrictions and capturing the repairs produced by speakers faced with phonotactically unattested sequences. This is accomplished through the interaction of RELEASE and ALIGN constraints. RELEASE constraints encode the phonetic environments that can give rise to acceptable gestural sequences, and temporal ALIGN constraints regulate how gestures are coordinated relative to one another.

The results of the acoustic experiment show that when asked to produce non-native word-initial consonant clusters, speakers reliably distinguish between clusters even though none of them are found in the legal English cluster inventory. An acoustic examination of the most common repair, vowel insertion, revealed that the inserted material was not the same as a lexical schwa. For the majority of speakers, the results of the ultrasound experiment were likewise not consistent with epenthesis, but rather with gestural mistiming. The production of the release, and

33

consequently a transitional vowel between the two consonants, provides a better perceptual or aerodynamic environment for the recoverability of the initial consonant. The interaction of RELEASE and ALIGN constraints also reflects these factors.

The Optimality Theoretic analysis developed in this paper accounts for the fact that speakers do not exhibit all-or-nothing performance on the non-native word-initial clusters, but rather produce them accurately some proportion of the time. It was proposed that speakers can treat constraints (here, ALIGN) as floating constraints, which can be reranked when speakers attempt to attain the grammar which allows such word-initial clusters. The idea that the final state of the native language grammar contains hidden rankings that can affect production under certain circumstances has important consequences for language acquisition, contact, and speech production.

Acknowledgements

References

Abrahamsson, N. (2003). Development and recoverability of L2 codas. *Studies in Second Language Acquisition, 25*(3), 313-349.

Anttila, A. (1997). Deriving variation from grammar. In F. Hinskens, R. v. Hout & W. L. Wetzel (Eds.), *Variation, Change, and Phonological Theory* (pp. 35-68). Amsterdam: John Benjamins.

Baayen, R. H. (2003). Probabilistic approaches to morphology. In R. Bod, J. Hay & S. Jannedy (Eds.), *Probabilistic Linguistics*. Cambridge: MIT Press.

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America, 57*, 456-461.

Blumstein, S., & Stevens, K. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64*, 1358-1368.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives.* The Hague: Holland Academic Graphics.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry, 32*, 45-86.

Boersma, P., & Levelt, C. (2000). Gradual constraint-ranking learning algorithm predicts acquisition order. *Proceedings of the Child Language Research Forum, 30*, 229-237.

Bradley, T. (2002). Gestural timing and derived environment effects in Norwegian Clusters. In L. Mikkelsen & C. Potts (Eds.), *Proceedings of WCCFL21*. Somerville, MA: Cascadilla Press.

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research, 7*(1), 35-59.

Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219-252.

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

Browman, C., & Goldstein, L. (1990b). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.

Browman, C., & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti & L. Raphael (Eds.), *Producing Speech: Contemporary Issues for Katherine Safford Harris*. New York: American Insititue of Physics.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Byrd, D. (1995). C-centers revisited. *Phonetica, 52*, 285-306.

Byrd, D. (1996a). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24*, 209-244.

Byrd, D. (1996b). A phase window framework for articulatory timing. *Phonology, 13*, 139-169.

Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Bloomington: Indiana University Press.

Cho, T. (1998). *Specification of Intergestural Timing and Gestural Overlap: EMA and EPG Studies*. Unpublished M.A. Thesis, UCLA.

Côté, M.-H. (2000). *Consonant Cluster Phonotactics: A Perceptual Approach*. Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics, 19*(6/7), 619-633.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics, 34*(1), 104-137.

Davidson, L., & Goldrick, M. (2003). Tense, agreement and defaults in child Catalan: An Optimality Theoretic analysis. In S. Montrul (Ed.), *Selected Papers from the 4th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Cambridge, MA: Cascadilla Press.

Davidson, L., Jusczyk, P., & Smolensky, P. (2004). The initial and final states: Theoretical implications and experimental explorations of richness of the base. In R. Kager, W. Zonneveld & J. Pater (Eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Davidson, L., & Legendre, G. (2003). Defaults and competition in the acquisition of functional categories in Catalan and French. In R. Nuñez-Cedeno, R. Cameron & L. López (Eds.), *A Romance Perspective on Language Knowledge and Use*. Philadelphia: John Benjamins.

Dorman, M., Studdert-Kennedy, M., & Raphael, L. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context dependent cues. *Perception & Psychophysics, 22*, 109-122.

Elson, B. (1956). Sierra Popoluca syllable structure. *International Journal of American Linguistics, 13*, 13-17.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory, 20*(2), 269-337.

Hall, N. (2003). *Gestures and Segments: Vowel Intrusion as Overlap.* Unpublished Ph.D.
dissertation, University of Massachusetts, Amherst, MA.

Hammond, M. (1999). *The Phonology of English*. Oxford: Oxford University Press.

Hansen, J. (2004). Developmental sequences in the acquisition of English L2 syllable codas.
Studies in Second Language Acquisition, 26, 85-124.

Henderson, J., & Repp, B. (1982). Is a stop consonant released when followed by another stop
consonant? *Phonetica, 39*, 71 - 82.

Kingston, J. (1985). *The Phonetics and Phonology of the Timing of Oral and Glottal Events.*
Unpublished Ph.D. dissertation, U.C. Berkeley.

Kingston, J. (1990). Articulatory binding. In J. Kingston & M. Beckman (Eds.), *Papers in
Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge:
Cambridge University Press.

Kirchner, R. (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry, 27*, 341-
350.

Kochetov, A. (2001/2002). *Production, Perception and Emergent Phonotactic Patterns: A Case
of Contrastive Palatalization*. New York: Routledge.

Kochetov, A. (to appear). Syllable position effects and gestural organization: Articulatory
evidence from Russian. In *Papers in Laboratory Phonology VIII*. Berlin: Mouton de
Gruyter.

Krakow, R. (1999). Physiological organization of syllables: a review. *Journal of Phonetics, 27*,
23-54.

Legendre, G., Hagstrom, P., Vainikka, A., & Todorova, M. (2001). Evidence for syntactic
competition in the acquisition of tense and agreement in child French. In A. Okrent & J.

P. Boyle (Eds.), *Papers from the 36th Regional Meeting of the Chicago Linguistics Society*. Chicago, Ill.: Chicago Linguistics Society.

Lubowicz, A. (1998). Derived environment effects in Optimality Theory. *Lingua, 112*, 243-280.

Matteson, E., & Pike, K. (1958). Non-phonemic transition vocoids in Piro (Arawak). *Miscellanea Phonetica, 3*, 22-30.

McCarthy, J., & Prince, A. (1993). Generalized alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology*. Dordrecht: Kluwer.

McCarthy, J., & Prince, A. (1999). Faithfulness and identity in Prosodic Morphology. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press.

McGowan, R. S., & Saltzman, E. (1995). Incorporating aerodynamic and laryngeal components into task dynamics. *Journal of Phonetics, 23*, 255-269.

Morelli, F. (1999). *The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory*. Unpublished Ph.D. dissertation, University of Maryland.

Moreton, E., & Smolensky, P. (2002). Typological consequences of local constraint conjunction. In L. Mikkelsen & C. Potts (Eds.), *Proceedings of the West Coast Conference on Formal Linguistics 21*. Cambridge, MA: Cascadilla Press.

Nagy, N., & Reynolds, W. (1997). Optimality Theory and variable word-final deletion in Faeter. *Language Variation and Change, 9*, 37-55.

Ohala, J. (1994). Speech aerodynamics. In R. E. Asher & J. M. Y. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (Vol. 8, pp. 4144-4148). New York: Pergamon Press.

Ohala, J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. H. Jahr (Eds.), *Language and Its Ecology*. Berlin: Mouton de Gruyter.

Padgett, J. (2001). *Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory*. ms, UC Santa Cruz.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge: MIT Press.

Reynolds, W. (1994). *Variation and Phonological Theory*. Unpublished Ph.D. dissertation, University of Pennsylvania.

Skousen, R. (1989). *Analogical Modeling of Language*. Dordrecht: Kluwer.

Smolensky, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science, 23*(4), 589-613.

Smolensky, P. (2005). Optimality in phonology II: Markedness, feature domains, and Local Constraint Conjunction. In P. Smolensky & G. Legendre (Eds.), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.

Smolensky, P. (2006). Harmony in linguistic cognition. *Cognitive Science*.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics, 21*, 291-311.

Steriade, D. (1993). Closure, release and nasal contours. In M. Huffman & R. Krakow (Eds.), *Nasals, Nasalization, and the Velum* (pp. 401-470). San Diego: Academic Press.

Steriade, D. (1997). *Phonetics in Phonology: The Case of Laryngeal Neutralization.* ms., UCLA.

Tesar, B., & Smolensky, P. (2000). *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.

Weinberger, S. (1994). Functional and phonetic constraints on second language phonology. In M. Yavas (Ed.), *First and Second Language Phonology* (pp. 283-302). San Diego: Singular Publishing Group.

Westbury, J. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America, 91*, 2903-2910.

Westbury, J., & Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*, 145-166.

Zsiga, E. (2003). Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition, 25*, 399-432.

Zuraw, K. (2000). *Patterned exceptions in phonology.* Unpublished Ph.D. dissertation, UCLA, Los Angeles.

Zuraw, K. (2003). Probability in language change. In R. Bod, J. Hay & S. Jannedy (Eds.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press.

Table 1. Possible response types

Response Type	Definition	Example
Correct	Target is produced with no changes or simplifications	/zgano/ → [zgano]
Insertion	Target is produced with a vocalic period between the consonants in the cluster	/zgano/ → [zᵊgano]
Deletion	Target is produced with either the first or second member deleted	/zgano/ → [zano] /zgano/ → [gano]
Prothesis	Target is produced with a vocalic period before the cluster	/zgano/ → [əzgano]
Segment Change	Target is produced with two segments, but one differs from the intended segment	/zgano/ → [sgano]
Other	Target is not produced, has more than one error, or is completely unrecognizable	/zgano/ → ∅ /zgano/ → [zəzano] /zgano/ → [skamu]

Table 2. Duration, first formant (F1) and second formant (F2) midpoint frequencies of lexical and inserted vowels.

		Duration (ms)	F1 (Hz)	F2 (Hz)
/fC/	Lexical	43.6	517	1596
	Inserted	31.3	456	1485
/zC/	Lexical	64.3	483	1741
	Inserted	45.0	423	1671
/vC/	Lexical	64.9	501	1596
	Inserted	42.5	427	1543

Table 3. Observed proportion correct for Speaker 8 versus predicted proportion correct based on uniform floating of ALIGN constraint.

	EXPERIMENT	THEORY	Uniform Floating	
Cluster Stratum	Observed Correct, Speaker 8	ALIGN Position	Visitation Probability	Predicted Correct
1. *legal clusters*	--	❶	25%	100%
2. *fm,fnfp,ft,fk,fs*	65%	②	25%	75%
3. *zm,zn,zb,zd,zg*	21%	③	25%	50%
4. *vm,vn,vb,vd,vg*	17%	④	25%	25%

Figure 1. Spectrograms illustrating different coordination patterns for word-medial stop clusters in English *laptop* (top) and Russian *vnaklatku* 'with sugar' (bottom). The important aspect of the spectrogram is the interval of the stop sequences [pt] and [tk]. In English, there is no release between the oral closure for the /p/ and the following /t/, meaning that the closure for /t/ begins before the constriction for /p/ is released. In Russian, the closure of the /t/ is released before the closure of the /k/ begins. The release of the first stop in a consonant sequence is signaled by a quick burst of energy that is visible on a spectrogram when the air built up behind the stop closure is expelled (the dark spike following the /t/ and preceding the closure for the /k/).
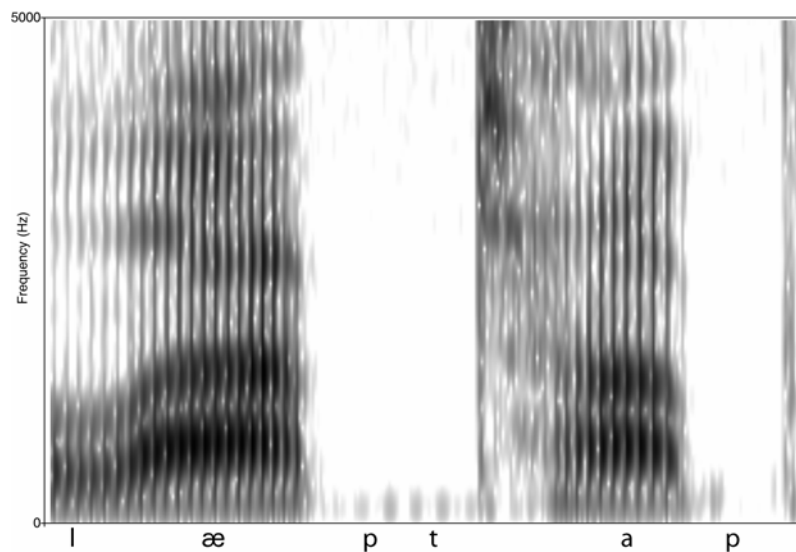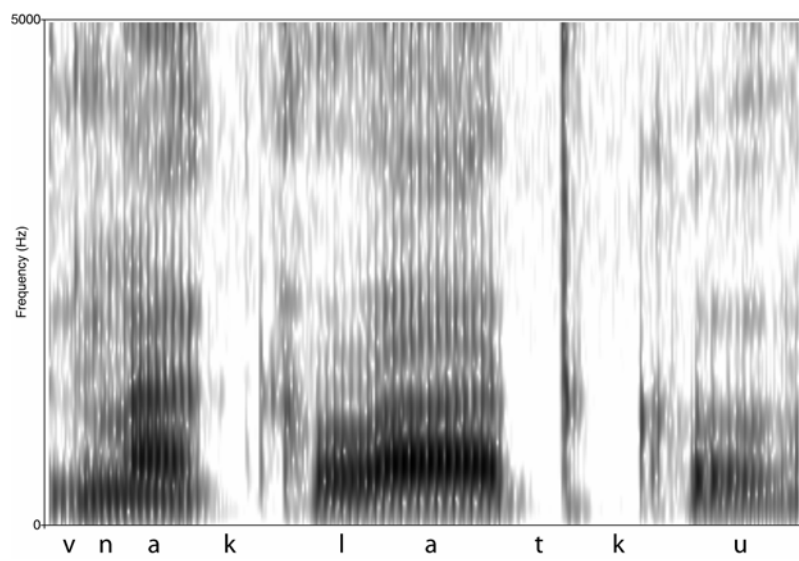
Figure 2. Spectrograms illustrating the production of the Slavic non-word *zgano* by a Czech speaker (top) and English speaker (middle). The bottom spectrogram is the English speaker producing /zəgano/ with an intentional schwa. Note that there is no voicing at all between the /z/ and the /g/ for the Slovak speaker, but that the English speaker has a period of voicing between /z/ and /g/. The duration of the /ə/ for the intended schwa in the bottom spectrogram is considerably longer than the duration of the transitional schwa in the middle panel.

z   ə   g   ɑ   n   o

z   ə   g   ɑ   n   o
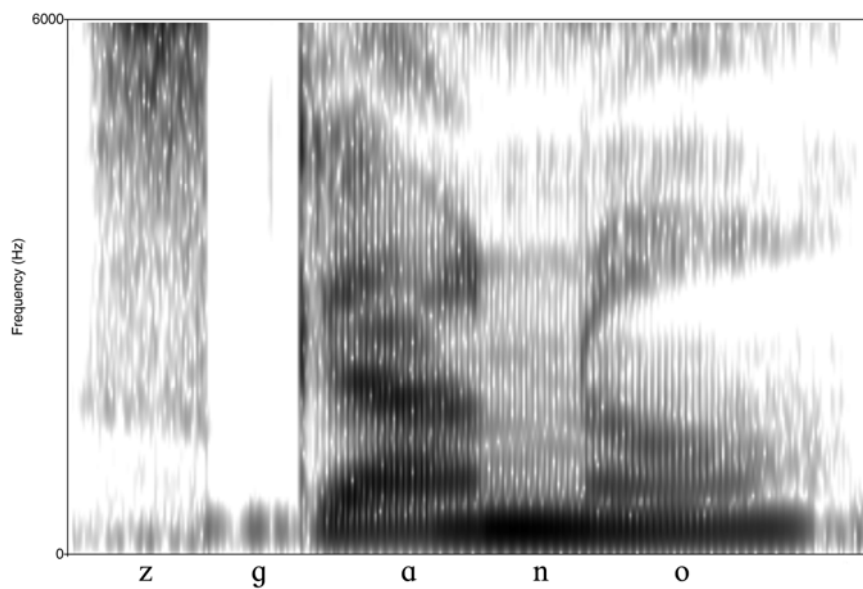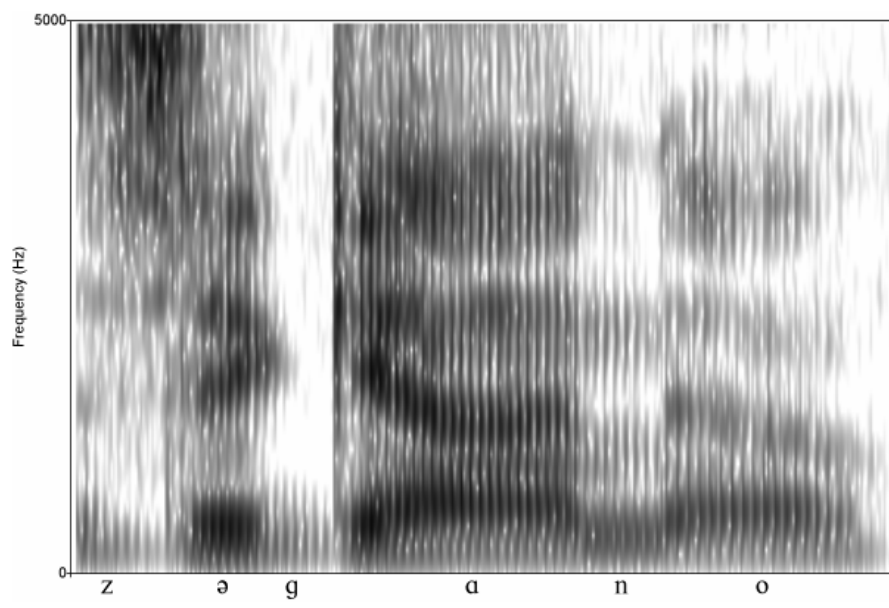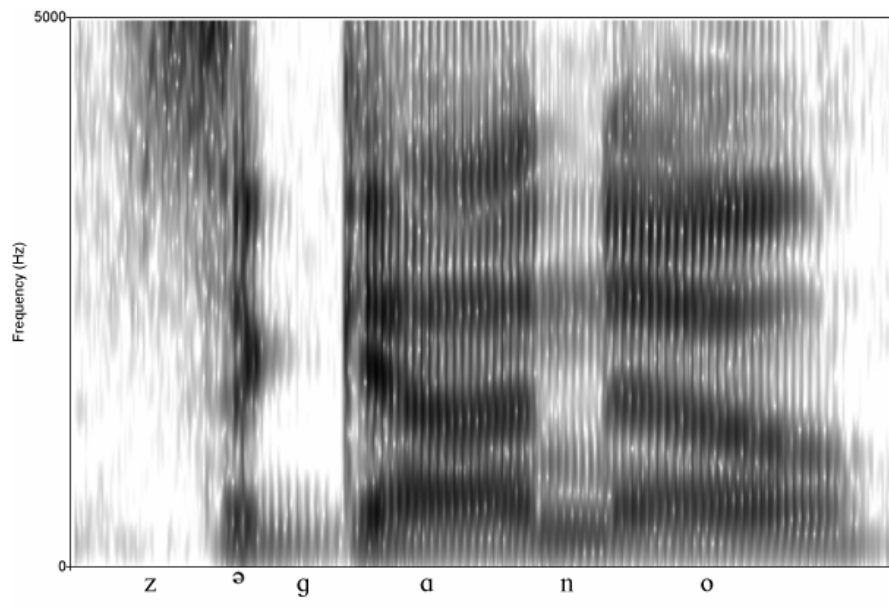
Figure 3. Gestural landmarks. The onset marks when the articulator begins moving toward the intended constriction, the target is when the constriction is achieved, and the release is when the articulator begins moving away from the constriction. The portion of the constriction between the target and the release is also referred to as the "plateau".
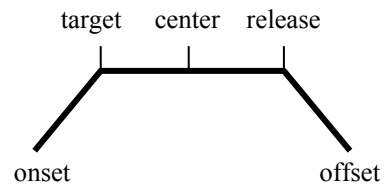
target   center   release

onset   offset

Figure 4. Close transition between consonants

C_1 C_2
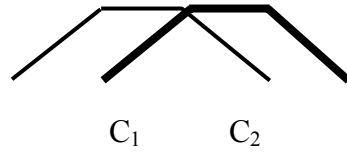
Figure 5. Open transition between consonants
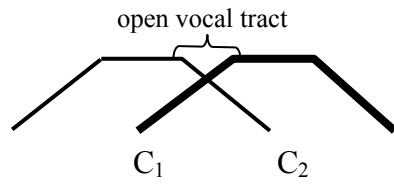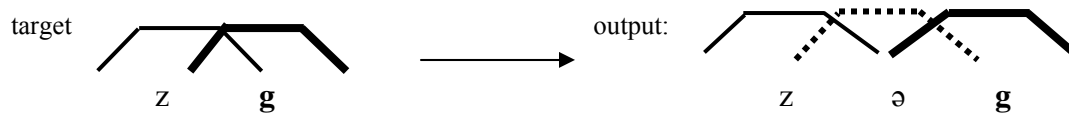

open vocal tract

C_1 C_2

Figure 6. Repair possibilities for speakers producing non-native sequences

a.   phonological epenthesis ("epenthesized schwa")

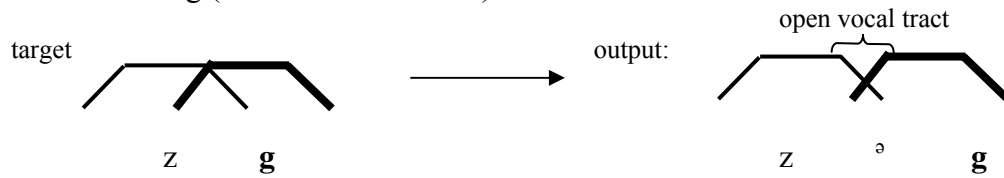b. gestural mistiming ("transitional schwa")

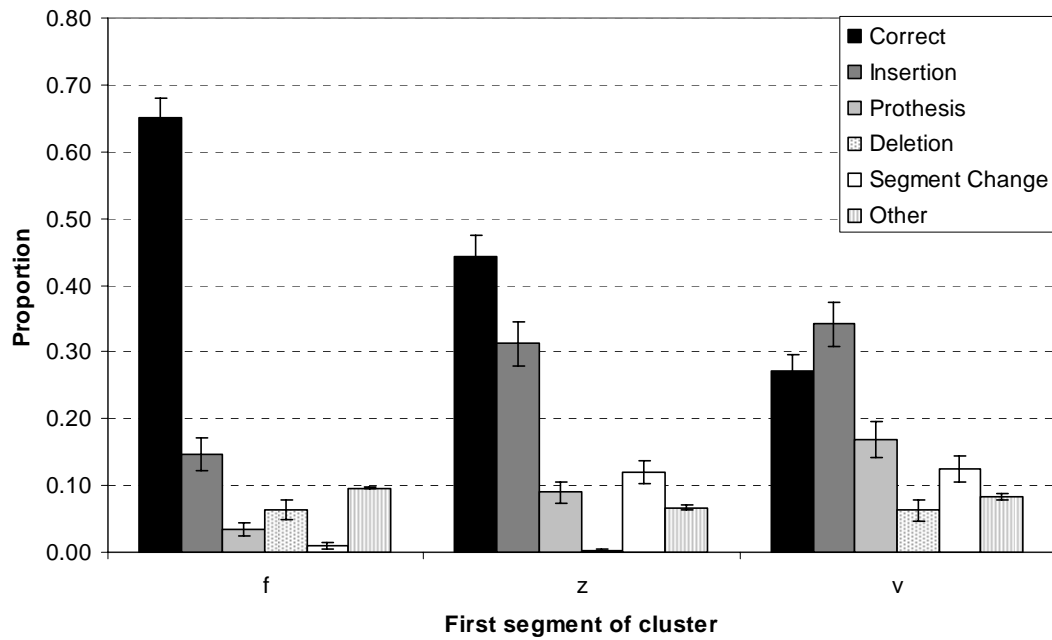Figure 7. Proportion of response type for each cluster type plotted by first segment. Error bars represent standard error.

Figure 8. Floating range for ALIGN

$$\begin{array}{rcl}
\ldots & \gg & ④ \\
\text{REL/[NoGroove],[NonAprx]\&REL/[VcdObs],[NonAprx]} & \gg & ③ \\
\text{REL/[VcdObs],[NonAprx]} & \gg & ② \\
\text{REL/[NoGroove],[NonAprx]} & \gg & ❶
\end{array}$$

ALIGN
(C_1, release,
C_2, target)