Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

External Sandhi in L2 Segmental Phonetics – Final (De)Voicing in Polish English

Geoffrey Schwartz
Adam Mickiewicz University

Anna Balas
Adam Mickiewicz University

Arkadiusz Rojczyk
University of Silesia

Abstract

The effects of external sandhi, phonological processes that span word boundaries, have been largely neglected in L2 speech research. The glottalization of word-initial vowels in Polish may act as a "sandhi blocker" that prevents the type of liaison across word boundaries that is common in English (e.g. *find out*/*fine doubt*). This reinforces the context for another process, final obstruent devoicing, which is typical of Polish-accented English. Clearly 'initial' and 'final' do not mean the same thing for the phonologies of the two languages. An adequate theory of phonological representation should be able to express these differences. This paper presents an acoustic study of the speech of voiced C#V sequences in Polish English. Results show that the acquisition of liaison, which entails suppression of the L1 vowel-initial glottalization process, contributes to the error-free production of final voiced obstruents, implying the internalization of cross-language differences in boundary representation.

Although research into the acquisition of second language (L2) speech has flourished in recent years, a number of areas remain to be explored. External sandhi, phonological processes that span word boundaries, constitute one such uncharted territory in L2 speech research. In what

follows we will briefly review some existing L2 sandhi research. Section 2 offers a representational perspective that makes predictions for further research into L2 sandhi acquisition. In Section 3 we present an acoustic study of Polish learners of English, which underlies the importance of sandhi acquisition for the production of word-final voiced obstruents in English.

Zsiga (2011) presents a literature review that is helpful in evaluating the current state of L2 sandhi research. She shows that this research has had a rather limited scope, focusing on a proposed Word Integrity (WI) constraint (Cebrian, 2000), by which learners are claimed to resist the transfer of L1 sandhi processes into an L2. Cebrian's study concentrates on a voicing assimilation process that occurs at word boundaries in Catalan, but does not frequently appear in Catalan English. Later studies (Zsiga, 2003; Lleo & Vogel, 2004) aimed at evaluating the WI hypothesis yielded some support for the proposal, but also encountered complications. In Zsiga's (2011) own study of Korean English, an L1 obstruent nasalization process was transferred to English with greater than expected frequency, challenging the WI hypothesis.

While the WI-oriented research has focused on L2 sandhi as a possible form of interference in L2, with the exception of liaison and enchaînement in L2 French, there is very little published research on the acquisition of L2 sandhi. This gap leaves a number of questions unaddressed. These questions concern the nature of WI itself. What is WI? Is it universal or language-specific? How can we characterize its effects for learners? With regard to this last question, WI has been seen on the one hand as a positive effect for learners in that it may inhibit L1 interference. On the other hand, what if WI is strong in L1 while the target language obscures word boundaries?

It is this situation that we claim to be observable in the speech of Polish learners of English. English is characterized by a large number of linking and liaison processes that obscure the boundaries between prosodic constituents. These processes often appear to take place at the syllable level, affecting consonants in coda position. For example, the phrase find out may be indistinguishable from fine doubt. The /d/ in find alters its syllabic affiliation to become an onset. The phrase got you may shorten to gotcha, in which the coda stop and the onset glide are fused into a single onset. In Polish, on the other hand, Word Integrity appears to be strong, as noted by Rubach & Booij (1990). A /t#j/ sequence (as in *kot jest* 'the cat is') never undergoes mutating palatalization that may occur at morpheme boundaries. Word-final codas are never resyllabified as onsets of vowel-

initial syllables. Thus, *byt Oli* 'Ola's existence' is distinct from *by Toli* 'so Tola (dat.)'. In addition, WI in Polish may be reinforced by a tendency to produce word-initial vowels with a glottal attack or a glottal stop (Dukiewicz & Sawicka 1995, Balas 2011, Schwartz 2013), so the phrase *Ewa i Ola* 'Ewa and Ola' may be heard as [ʔeva ʔi ʔola]. In sum, L1 Polish appears to block the types of sandhi processes that reorganize prosodic constituents in languages such as French and English.

Zsiga (2011) notes that external sandhi research can inform us about the interaction of grammar and the lexicon, as well as phrasing and speech planning. External sandhi may also shed light on phonological representations in L1 and L2. We suggest that adequate descriptions of these representations must tell us something about the nature of boundaries themselves, an issue we turn to in the following section.

THE PHONOLOGY OF BOUNDARIES IN POLISH AND ENGLISH

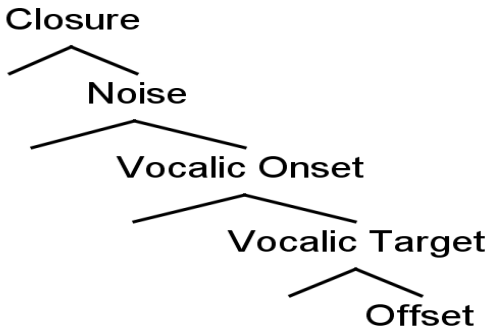
The relative paucity of L2 sandhi research may stem from theoretical difficulties concerning the representation of prosodic boundaries. As one might expect, L2 speech research is frequently based on relatively uncontroversial phonological parameters such as vowel quality and Voice Onset Time (VOT). By contrast, while describing phonological behavior at prosodic boundaries may be straightforward, there is little consensus as to what those boundaries are made of and how they arise. It is therefore reasonable to ask the following questions. What does it mean to say that a sound is 'final' or 'initial'? Are these terms equivalent across languages?

Traditionally, boundaries have been represented by diacritic marks such as # (e.g. Chomsky & Halle, 1968), or bracketed domains in a hierarchy of prosodic constituents at the syllable level and above (e.g. Nespor & Vogel, 1986). Scheer (2008) argues that while such strategies are convenient to describe where various sandhi processes may or may not apply, they are essentially arbitrary in nature. Likewise the terms 'initial' and 'final' are meaningless if they make reference to such diacritic symbols. Consider the English phrase *find out*. In native English the /d/ is generally pronounced as initial – the phrase is usually indistinguishable from *fine doubt*. By contrast, in a similar sequence in Polish, the /d/ is clearly 'final', while the following vowel is typically glottalized (Dukiewicz & Sawicka, 1995), underlying its 'initial' status. An adequate model of phonological representation should be able to capture these cross-linguistic differences.

The Onset Prominence framework (OP; Schwartz, 2013) contains phonetically-derived representational parameters that may explain language-specific differences in sandhi processes. The crucial aspect of the model is that prosodic constituents and segmental representations are constructed from the same materials, so boundaries fall out directly from the standard phonotactic mechanisms of the framework.

Before discussing the specifics of boundary formation, it is necessary to provide some background discussion about the primitives employed in the OP environment. In the model, segments are derived from a primitive prosodic constituent shown in (1), which represents a hierarchy of phonetic events associated with a stop-vowel sequence. A stop-vowel sequence will contain all four layers of the structure in (1). Place and laryngeal specifications attach to the terminal nodes.

(1) The Onset Prominence representational environment (Schwartz, 2013)



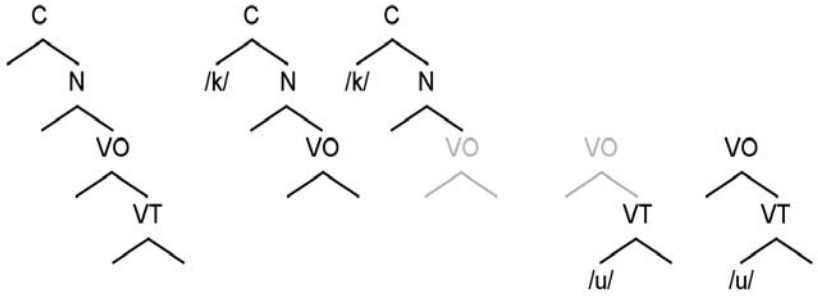
In a CV sequence the initial portion of vowels, represented by the Vocalic Onset (VO) node in (1), contains an inherent ambiguity with regard to the traditional categories of 'consonant' and 'vowel'. Acoustically, this transition is characterized by the high amplitude periodicity and robust formant structure commonly associated with vowels. Yet listeners rely on acoustic cues at vowel onset for the identification of the previous consonant (e.g. Wright, 2004). This ambiguity captures an important concept from the speech perception literature: the failure of speech to meet the 'linearity condition' (e.g. Wright et al., 1997), by which any given portion of the acoustic signal is presumed to contain cues to one and only one segment. Vowel onset contains acoustic cues to the identity of the vowel itself as well as the preceding consonant. It is a point of overlap between 'segmental' units.

The OP environment encodes this ambiguity by splitting vowels into two structural layers: Vocalic Onset (VO) and Vocalic Target (VT). This configuration creates representational parameters, shown in (2), for the

ambiguity associated with VO. In (2), we see on the left the primitive CV constituent from which both consonants and vowels are derived. The trees in the centre illustrate how VO may or may not be contained in higher-level obstruents. Finally, on the right we see that VO may or may not dominate lower level VT nodes in the representation of vowels. The fading in these structures may be interpreted as a placeholder, defining the level of neighboring nodes within the OP hierarchy.

From the representational perspective adopted here, we may replace the traditional ONSET constraint with a well-formedness condition on prosodic consituents; a minimal 'syllable' must contain active structural nodes both below and above the VT level. On this view, the rightmost tree, a vowel /u/ containing a VO node, is a well-formed constituent that contains an element that is analogous to an 'empty consonant' (Marlett & Stemberger, 1983). OP representations, however, show that this node is not empty. It is an essential building block of prosodic structure that may be strengthened by means of gliding, glottalization, or consonant epenthesis. Our claim for Polish is that this node is built into the representation of vowels, which are often realized with a glottal attack word-initially.

(2) VO parameters in the representation of consonants and vowels

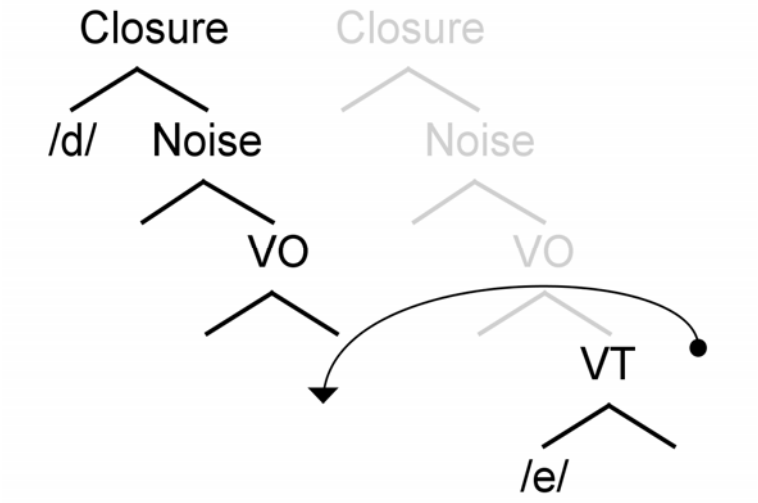


In the OP environment, the formation of boundaries results from the same principles by which segmental structures combine to form 'syllables'. The most fundamental of these mechanisms is the *absorption* of the lower-level vowel structures into the higher-level obstruents to form a CV.[1] This is depicted in (3), in which a vowel /e/ is absorbed into a stop /d/, which would result in a CV 'syllable' /de/. As pointed out earlier, this

[1] In the OP environment, as in other representational theories, 'codas' such as the final /k/ in *quick* are derived from the 'onset' status of consonants. The position of the final /k/ in *quick* is the result of a submersion process that will not be relevant for the present paper.

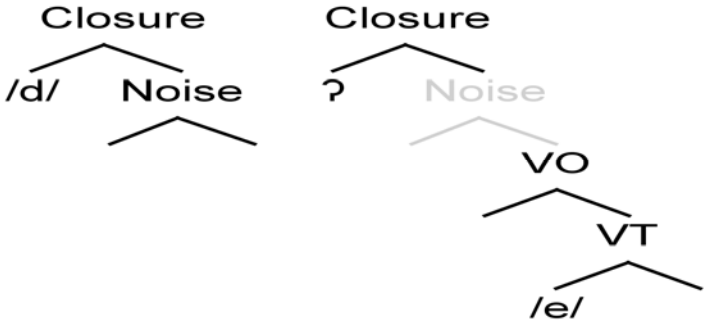
process is common at orthographic boundaries in English – in such cases there is often no prosodic boundary.

(3) Absorption of vowel into preceding stop structure



By contrast, a vowel that is specified with the VO node is structurally stronger, and the realization of this node is susceptible to fortification by means of glottalization. This is expressed in the OP environment as the strengthening of a VO-specified vowel, achieved by the activation of a Closure node as we see in (4). Absorption of a Closure-specified structure is not possible and a boundary is formed. Note in (4) that there is no need to stipulate that the /d/ is 'final' and the /e/ is 'initial'; the boundary arises when two adjacent segmental structures cannot merge.

(4) Absorption of vowel into stop blocked by glottalization



It appears as though Polish and English differ with regard to the representation of vowels following orthographic boundaries. In Polish, a larger vocalic structure containing VO is preserved through glottalization. This in turn reinforces the 'final' status of the preceding consonant,

opening the door to a well-documented element of Polish-accented English: final obstruent devoicing. The next section will examine the interaction of glottalization and final devoicing by means of an acoustic phonetic study.

EXPERIMENTAL METHOD

This section will describe an acoustic study on English C#V (where the consonant is underlyingly voiced) sequences in the speech of Polish learners[2]. The representational parameters of the OP model allow us to make predictions regarding the interaction between the realization of word-initial vowels (with or without glottalization) and the production of the preceding word-final consonant. The basic hypothesis that we are testing is that the glottalization of word-initial vowels, a process that appears to be a robust element of Polish phonology (Dukiewicz & Sawicka, 1995; Schwartz, 2013), is a form of L1 interference in Polish English. Glottalization reinforces the 'final' status of the consonant in C#V sequences and thereby contributes to final devoicing.

Subjects. Our analysis is based on recordings of 16 Polish students majoring in English studies at the University of Silesia, Sosnowiec (n=6) and Adam Mickiewicz University, Poznań (n=10). Since our focus in this study is on issues of representation, we will not be making direct comparisons of various years of study or levels of proficiency. Future work will look at learners at the secondary school level and investigate differences among university students.

Materials. The linguistic materials were comprised of a sentence list in English containing 35 C#V sequences in which the consonant was a voiced obstruent. The recordings were made in a sound-proof recording booth at each of the two universities. The chambers were equipped with studio-quality audio interfaces allowing for high quality recordings directly onto a laptop computer. The sentences were presented to the subjects in the form of a PowerPoint presentation. The subjects performed two tasks. In the first task, students simply read the sentences that were shown on the PowerPoint slide. In the second task, the slides were accompanied by audio recordings of a native speaker of English producing the given

[2] This study follows up on a pilot experiment described in Schwartz (2012), and covers a greater number of speakers and consonantal contexts. It also provides additional measures of consonant voicing, and more precise coding of glottal events.

sentence, and students were instructed to repeat after the recordings that they heard. The total number of tokens for analysis was 1120 (16 speakers * 2 tasks * 35 tokens).

ACOUSTIC ANALYSIS

Acoustic analysis was performed by hand with the help of the Praat program (Boersma & Weenink 2011), and focused on both voicing parameters associated with the final consonant as well as the realization of the vowel. With respect to voicing, the following measurements were made.

- Duration of the preceding vowel (in milliseconds)
- Duration of the consonant (in milliseconds; closure+noise in the case of stops, noise in the case of fricatives)
- Duration of periodicity (vocal fold vibration) during the consonant articulation (in milliseconds).

From these measurements, voicing of the consonant was quantified according to two calculated variables.

- V/C ratio: duration of the preceding vowel divided by duration of the preceding consonant. V/C ratios are higher when the consonant is voiced (Chen, 1970), an effect which is especially dramatic and perceptually relevant in English (Port & Dalby, 1982)
- %Voiced: the duration of the voiced period of the consonant divided by the total consonant duration multiplied by 100. This measure allows us to describe how much of a given consonant was in fact voiced.

With regard to the realization of the vowel in the C#V sequences, the goal of the acoustic examination was to determine the presence or absence of glottalization according to guidelines laid out in the phonetics literature (e.g. Dilley et al., 1996). Preliminary analysis allowed us to establish three categories. When no glottalization (or pause) was visible on the vowel, the tokens were labeled as 'liaised' (L). Tokens that showed glottal stops or some degree of glottalization on the vowel were described as 'glottalized and unliased' (labelled GU). Finally, in some cases glottalization was visible on the vowel, but not immediately. These tokens appeared to contain an 'intrusive vowel' (labelled IV), in which a short vocoid is

visible between the final consonant and the glottal event. An example of an IV token is given in Figure 1, which shows one subject's production of the C#V sequence in *showed everyone*. The intrusive vowel preceding the glottal stop is selected in the waveform/spectrogram display.

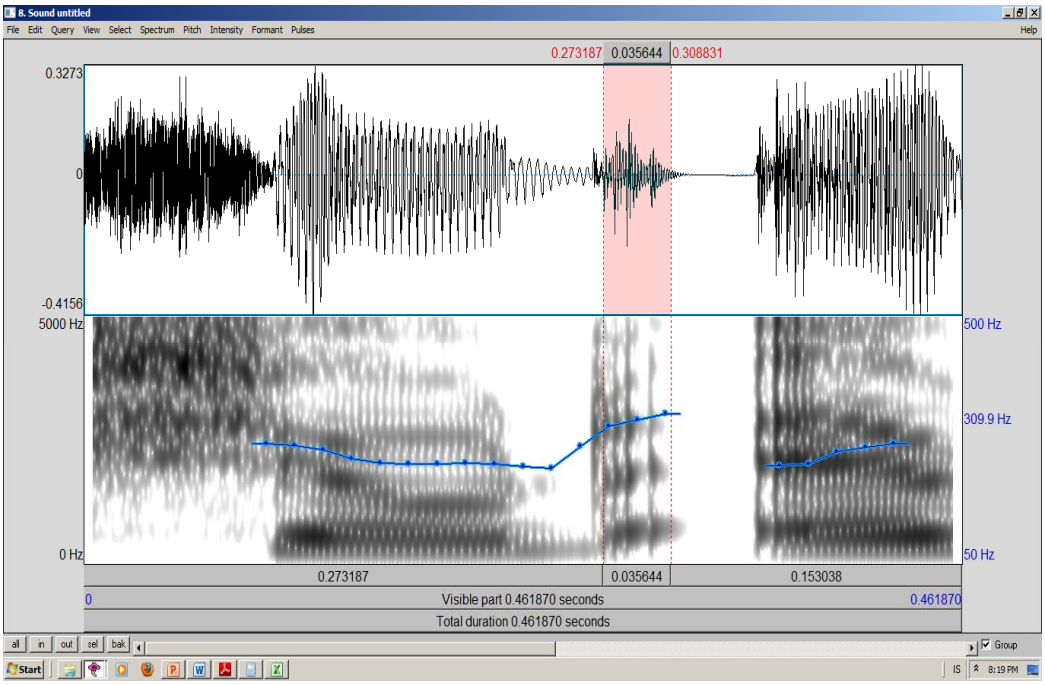


Figure 1. Waveform and spectrogram of the C#V sequence in *showed everyone*. Selection indicates the intrusive vowel.

RESULTS

The final results for glottalization type were as follows. Glottalized and unliaised tokens (GU) were found in 46.2 % of the items, 31% were liaised (L), and 22.8% were realized with an intrusive vocoid (IV). To measure the effects of glottalization on the voicing parameters, one-way ANOVAs were performed using the voicing parameters as dependent variables. The results of the analysis for V/C ratio are shown in Figure 2. GU tokens showed a significantly lower V/C ratio than the other realizations of the initial vowels. The difference between L and IV tokens was non-significant [$F(2,1097)=95.86$, $p<.001$; Post-Hoc Bonferroni: GU-L: $p<.001$; GU-IV: $p<.001$; L-IV: $p=ns$].

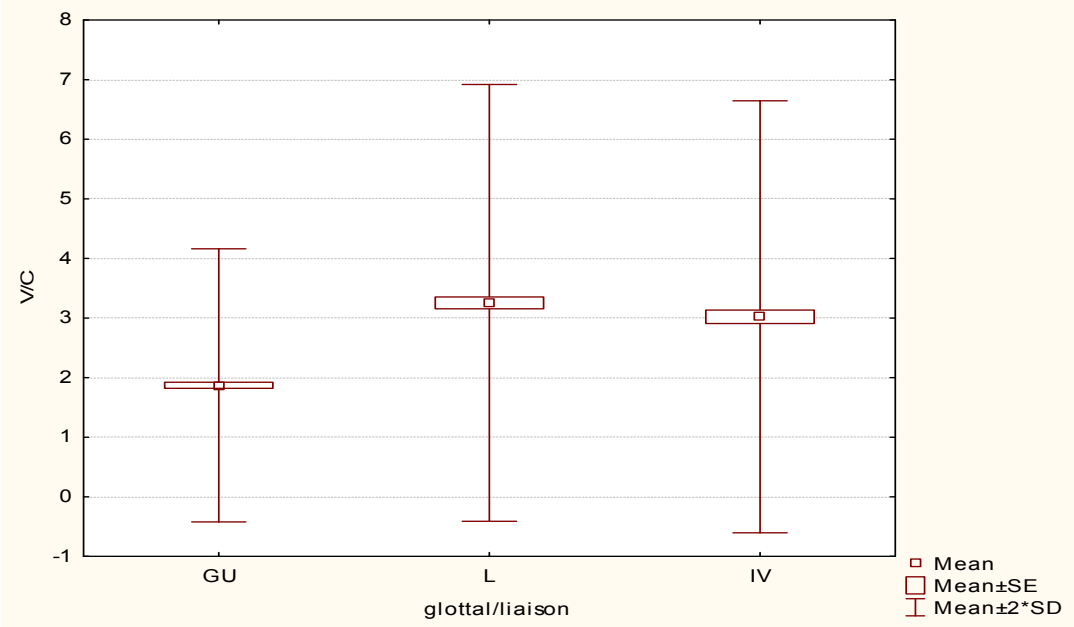


Figure 2. Boxplots representing V/C ratios for the three glottalization types

The results for the %Voiced parameter are given in Figure 3. Liaised tokens showed the most voicing, followed by IV tokens, and GU tokens with the least voicing. There was a main effect of glottalization type on %Voiced ratios and Post-hoc tests showed that all the pairwise differences were significant: [(F: 2,1097)=374,22, p<.000; Post Hoc Bonferroni: GU-L: p<.001; GU-IV: p<.001; L-IV: p<.001].

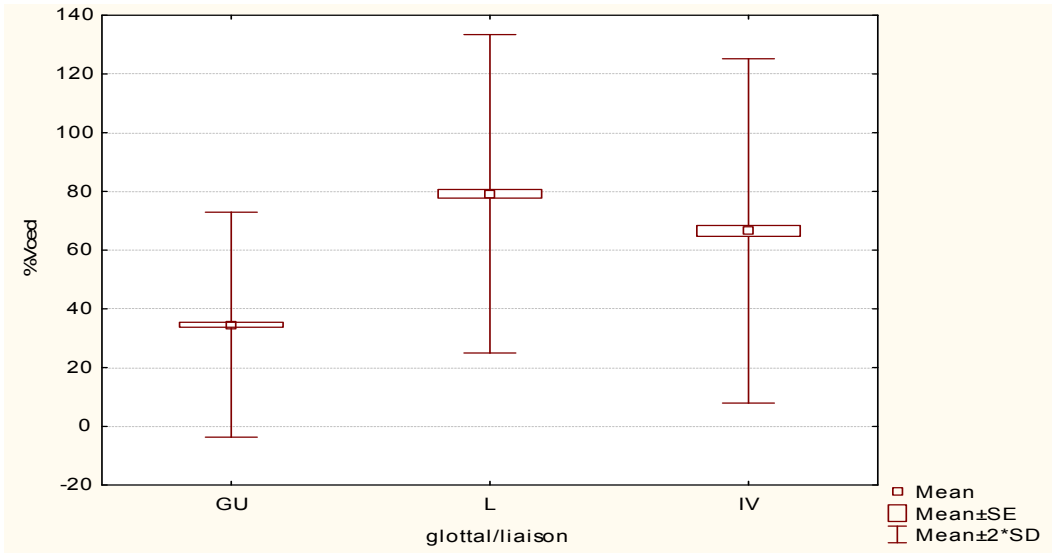


Figure 3. Boxplots representing %Voiced ratios for the three glottalization types

DISCUSSION AND LIMITATIONS

The results of our study reveal that glottalization of initial vowels is a clear predictor of final devoicing in the speech of Polish learners of English. As a consequence, it may be argued that C#V linking or liaison, a boundary effect that is predicted by the OP model for English, may have important implications for the realization of individual segments in L2 English speech.

Examining our study in further detail, it is interesting to compare the liaised tokens to those items that showed an intrusive vowel. Assuming that final voiced obstruents present a universal phonetic challenge for learners and native speakers alike, both liaison and vowel intrusion may be seen as strategies for maintaining voicing. In each case, speakers tamper with the 'final' status of the consonant. As suggested by English pronunciation textbooks, liaison appears to be a strategy employed frequently by native speakers. This was also reflected in the native models used for the imitation task in our study.

Figure 4 summarizes the task effects on both voicing and glottalization parameters, including comparison with the native baseline items. In the figure we can observe that liaison dominates in the native model. Of the 35 baseline tokens, 27 of them were produced with liaison, and only 6 were produced with an intrusive vowel (2 items were labeled GU). The other important trend to notice in Figure 4 concerns the relative likelihood of vowel intrusion and liaison in the two tasks. When final voiced obstruents were produced in the reading task, voicing tended to be facilitated by vowel intrusion rather than the native strategy of liaison. However, the imitation results suggest that native-style linking is indeed within reach of learners. Future work should address the long-term effects of the type of imitation tasks employed here.

In sum, the results of our study emphasize the importance of sandhi acquisition in the production of native-like L2 speech. The effects of L2 sandhi may be expected to be reflected in listener ratings of 'accentedness' of L2 speech. At the same time, it is entirely possible that with regard to boundary phenomena in English, accentedness may be inversely correlated with comprehensibility and/or intelligibility. In other words, native processes like liaison that obscure prosodic boundaries may be expected to be especially challenging for non-native listeners. Further experiments are underway to assess the perceptual implications of the acoustic data described here.

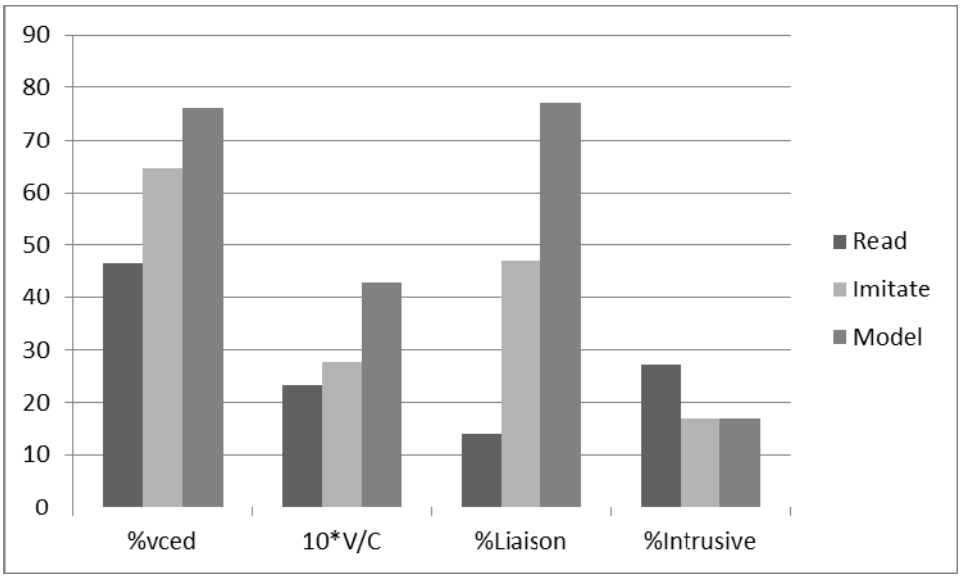


Figure 4. Voicing and glottalization parameters across the three tasks

ACKNOWLEDGEMENTS

Thanks to Ron Kim, Dave Leonard, and of course all of the participants in our experiment. This research was supported by a grant from the Polish National Science Centre (Narodowe Centrum Nauki), project nr.UMO-2012/05/HS2/04036.

REFERENCES

Balas, A. (2011). Glottal stops produced by Polish native speakers in Polish and in English. *Proceedings of ICPhS XVII,* 280-283.

Boersma, P. & Weenink, D. (2011). Praat: Doing phonetics by computer. [Computer program]. Version 5.2.18. Date of access: 22.03.2011.

Cebrian, J. (2000). Transferability and productivity of L1 rules in Catalan-English interlanguage. *Studies in Second Language Acquisition, 22,* 1-26.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22,* 129-159.

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

Dilley, L., Shattuck-Hufnagel, S. & Ostendorf, M.. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24,* 423-444.

Dukiewicz, L. & Sawicka, I. (1995). *Gramatyka współczesnego języka polskiego – fonetyka i fonologia* [Grammar of modern Polish – phonetics and phonology]. Kraków: Wydawnictwo Instytutu Języka Polskiego PAN.

LLéo, C. & Vogel, I. (2004). Learning new segments and reducing domains in German L2 phonology: The role of the Prosodic Hierarchy. *International Journal of Bilingualism, 8,* 79-104.

Marlett, S. & Stemberger, J. (1983). Empty consonants in Seri. *Linguistic Inquiry, 14,* 617-639.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Port, R. & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Attention, Perception and Psychophysics, 32,* 141-152.

Redi, L. & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29,* 407-429.

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

Scheer, T. (2008). Why the Prosodic Hierarchy is a diacritic and why the Interface must be Direct. In Hartmann, Jutta, Veronika Hegedüs & Henk van Riemsdijk, (eds.). *Sounds of Silence: Empty Elements in Syntax and Phonology*, 145-192. Amsterdam: Elsevier.

Schwartz, G. (2012). Initial glottalization and final devoicing in Polish English. *Research in Language, 10,* 159-171.

Schwartz, G. (2013). A representational parameter for onsetless syllables. *Journal of Linguistics 49*(3), 613-646.

Wright, R., Frisch, S., & Pisoni, D. (1997). Speech Perception. Research on spoken language processing, Progress Report No. 21. Ms., Indiana University.

Wright, R. (2004). Perceptual cue robustness and phonotactic constraints. In Hayes, B., R. Kirchner and D. Steriade (eds). *Phonetically Based Phonology.* Cambridge: Cambridge University Press, 34-57.

Zsiga, E. (2003). Relearning consonant timing. *Studies in Second Language Acquisition, 25,* 399-432.

Zsiga, E. (2011). External *sandhi* in a second language: The phonetics and phonology of obstruent nasalization in Korean and Korean-accented English. *Language,* 87, 289 – 345.