Phonemes, features and allophones in L2 phonology

Polish sibilants in Croatian ears (and brains)

Małgorzata Ćavar, University of Zadar & Silke Hamann, Heinrich-Heine Universität, Düsseldorf

Abstract

This article presents the results of a perception study of foreign phonemes, in particular, Polish phonemes categorized by Croatian native listeners. Our experiment shows that listeners can identify phonemes of an unknown/new language (L2) without having had prior training in this new language if the phonemes utilize the same phonological distinctive features as employed in the native language (L1). This extends to L2 sounds that have a different combination of features than the L1 phonemes. The results indicate that subjects use their L1 phonological knowledge in extrapolating distinctions in L2 sounds and that phonological features seem to play a larger role in L2 perception than phonemes.

1. Introduction

The branch of second language (L2) phonology is still a largely understudied branch of linguistics. This is probably due to the fact that L2 phonology involves two aspects: the acquisition of abstract, phonological categories (such as phonemes and features) and also the acquisition of the phonetic realizations of these L2 sounds (both the perceptual contrasts and the particular articulations that correspond to these percepts). Both, the categories and the phonetic realizations can be identical, overlapping or totally different from those used in the L1. In the acquisition process, L2 learners start with learning to perceive the relevant perceptual cues in the new language and to correlate these with abstract phonological representations. The mastery of the articulation of the sounds occurs only later. This observed acquisition order of first learning to map the percept to abstract representations and then learning to articulate these abstract representations seems to be incompatible with most phonological models in use, which describe the production process only, and totally neglect the perception process (but see Boersma 1998 or Stampe 1979 and Donegan 2001 for notable exceptions). Such production-only phonological approaches potentially led to the systematic misanalysis of L2 phonological effects.

At the same time we find a large amount of experimental studies on L2 perception. These, however, often lack an explicit phonological approach to test. Especially with respect to the abstract units relevant for perception, namely phonological features and phonological segments, we find only little work that aims at supporting or rejecting the relevance of one of the two. Most researchers working in experimental approaches seem to be in favor of the phoneme. For example, the Perceptual Assimilation Model (Best, McRoberts and Sithole 1988), which assumes that the more similar sounds of L1 and L2 are, the easier L2 sounds are acquired, is coined in terms of phonemes. Similarly, the Speech Learning Model (Flege 1987, Aoyama et al. 2004, etc.) views the learning process as allophone-by-allophone learning. This approach predicts no generalization across different phonemes, at least not at an early stage of the acquisition of phonology. More linguistically oriented approaches, on the other hand, favor the feature approach. Brannen (2002), for example, explains patterns of phone substitutions (in particular for English interdentals) in L2 in terms of the featural make-up of the L2 sounds. In the same vein, Eckman and Iverson (1994) account in terms of features for repair strategies in the production of L2 structures that are disallowed in L1.
Few studies make an attempt to disambiguate the issue and to provide evidence either for the phoneme or for the feature. Among them, Strange et al. (1998) indirectly supports the phoneme. By studying English vowel contrasts as perceived by Japanese learners, they discovered few parallels in the acquisition of the perception of vowels sharing the same features. On the other hand, Silbert et al. (2005) and de Jong & Silbert (2006) studied perception patterns by Korean L2 learners of English. They found that voicing generalizes across places and manner of articulation but L2 manner of articulation does not. Instead, L2 manner of articulation neutralizes in the L1 positions of neutralization. Thus, their results are so far inconclusive and experiments need to be extended.

In the present study, we focus on these phonological units of perception and describe an experiment that tested the relevance of the phonological feature for L2 speech perception. In our study, we expose naïve subjects to foreign phonemes. Our assumption is that the amount of misperceptions in the experiment depends on the phonological system of the native language of the subjects. In particular, we follow the assumption of de Jong and colleagues (p.c., also e.g. de Jong and Silbert 2006): if phonemes that share phonological features are misperceived equally often, this indicates that the relevant unit in L2 phonological perception is the phonological feature. If such phonemes, on the other hand, show differing results in a perception task, the phonological feature seems not relevant in L2 perception.

The article is organized in the following way. In section 2, we describe the sibilant systems of the two languages that we use in our perception experiment. In section 3 the experimental setting is described. Section 4 provides a discussion and section 5 the conclusions.

2. The non-anterior sibilants in Polish and Croatian

In the present study we look at the perception of Polish sounds by Croatian naïve listeners, focusing on the post-alveolar sibilants. Polish and Croatian have similar but not identical post-alveolar sibilant inventories. Both distinguish two places of articulation that we refer to as alveolopalatal and retroflex (for a discussion why the latter can be referred to as retroflex, see Hamann 2004).

Polish has a symmetrical post-alveolar sibilant system with two affricates and two fricatives at both places of articulation, namely the alveolopalatal [tɕ, dʑ, ɕ, ʑ] and the retroflex [tʂ, dʐ, ʂ, ʐ]. In terms of phonological features, post-alveolar sounds are all specified as [CORONAL, –anterior]. A possible feature specification that distinguishes between the Polish non-anterior sibilants is given in Table 1 (cf. Rubach 1984):

Table 1: Polish non-anterior coronal sibilants and their feature specification

	tʂ	**dʐ**	**ʂ**	**ʐ**	**tɕ**	**dʑ**	**ɕ**	**ʑ**
[continuant]	–	–	+	+	–	–	+	+
[voice]	–	+	–	+	–	+	–	+
[back]	+	+	+	+	–	–	–	–

Croatian also has a distinction between alveolopalatal and retroflex post-alveolar sibilants, but only for affricates. The alveolopalatal fricatives are lacking. The Croatian inventory of [CORONAL, –anterior] sibilants and their further feature-specification are presented in Table 2 (cf. Brozović 1972):

Table 2: Croatian non-anterior sibilants and their feature specification

	tʂ	**dʐ**	**ʂ**	**ʐ**	**tɕ**	**dʑ**
[continuant]	–	–	+	+	–	–
[voice]	–	+	–	+	–	+
[back]	+	+	+	+	–	–

These two sibilant systems differ only in terms of phonemes, Croatian has two phonemes less, but not in terms of features: both have to employ a feature like [continuant] to distinguish affricates from fricatives, a voicing feature, and a feature like [back] to distinguish the two post-alveolar places of articulation. They therefore provide an ideal testing ground for the question whether phonemes or features are of relevance in the perception of speech sounds: if we assume that native phonemes alone are used in speech perception, then we should find difficulties (more mistakes) for Croatian listeners when asked to categorize the Polish sounds [ɕ, ʑ]. If, on the other hand, the native features can be employed in such a task, then the Croatians should perform similar to Polish listeners, because [ɕ, ʑ] are composed of features known to Croatian listeners.

Phonetically, the sounds [ɕ, ʑ] are not absent from Croatian. [ɕ] and [ʑ] are allophones of /ʂ/ and /ʐ/, respectively, when in the context of prepalatal affricates, as in (1):

(1)		spelling	gloss
	gro[ʑdʑ]e	grožđe	‘grapes’
	li[ɕtɕ]e	lišće	‘leaves’

As [ɕ] and [ʂ] are allophones of one underlying phoneme, listeners are ‘trained’ to overlook the fine phonetic distinction between the two. The same should be valid for Croatian [ʑ] and [ʐ]. Guion and Lee (2006:123) remark that “[...] acoustic information not used to cue phonological distinctions comes to be systematically under-attended.” If we assume the phoneme as main perceptual unit, the fact that [ɕ, ʑ] appear in Croatian as allophones should contribute additionally to the difficulty of the task and result in a worse identification ratio for fricatives than for affricates. If we assume the feature, however, the presence of the allophones [ɕ, ʑ] should not influence the performance of Croatian listeners.

A further issue we have to take into consideration here are the phonetic realizations of the place contrast in the regional variants of Croatian. Standard Croatian prescribes the preservation of this contrast in spelling and pronunciation; however, the regional variants of the Croatian language differ strongly in this respect. We summarize here the three main dialect areas of Štokavian, Kaikavian, and Čakavian: Kaikavian is spoken in Zagorje, i.e. the area around Zagreb, Čakavian on the islands and the narrow strip of the mainland along the coast, and Štokavian in the rest of the country, as well as in Bosnia and Hercegovina and Serbia (Brozović 1970, Lončarić 1996). In the Štokavian of some areas of e.g. Bosnia and Hercegovina, the realization of the two sibilant places is virtually identical to the realization of the contrast in Polish. Other areas have at least partial neutralization, e.g. the Štokavian of Slavonia, and Kaikavian. In other areas, the contrast has shifted, for example, in the area of Zadar the alveolopalatals are realized as secondarily palatalized alveoalars, i.e. [t^J, d^J].

Croatians are trained to preserve the contrast from the first class of the primary school on, however, it seems that the training restrains itself to teaching the “correct” spelling and does not go that far as to teach the ‘correct’ pronunciation.

3. Categorization experiment

3.1 Speech material

As stimuli we used the Polish post-alveolar sibilants ([tɕ, dʑ, ɕ, ʑ, tʂ, dʐ, ʂ, ʐ]). These appeared in three different positions, namely, word-initially followed by a vowel, medially between two identical vowels and word-finally after a vowel. The vowels were either [a] or [ɛ]. This resulted in 48 items, none of them were real words in Polish or Croatian. By this we avoided lexical familiarity effects that might influence the categorization of the sounds in question.

The stimuli were produced by a female Polish native speaker at a normal speaking rate. They were recorded on a Marantz PMD 620 digital audio recorder at a sampling rate of 48 kHz in a quiet surrounding.

3.2 Subjects

Twenty subjects, all native-speakers of Croatian without any diagnosed hearing or speech impairment, have been tested. Their age ranges between 19 and 30 years, all of them are students or faculty members of the English department at the University of Zadar. None has learned Polish (or any other language with a similar sibilant inventory, for instance, Mandarin).

In addition, we had three control groups. The first group was two Polish native speakers, aged 31 and 34, one with and one without any linguistic background. The Polish control group was chosen to see whether the categorization experiment could be performed at all, and to have maximum correct identification scores to which we can compare the performance of the other groups.

The second group was three Slovenian subjects, aged 19 to 34, students and faculty members of the Slovenian Studies Department at the University of Nova Gorica; two subjects learnt Serbo-Croatian at school, one does not know Serbo-Croatian at all. The third group was three German native speakers, aged 20 to 24, all students at the linguistics department of the University of Düsseldorf. The German students had followed an introduction in

phonetics. Slovenian and German were chosen because both of these languages have only one series of non-anterior sibilants, namely the post-alveolar [ʃ ʒ] in German and the post-alveolar [ʃ, ʒ, tʃ, dʒ] in Slovenian.

None of the subjects in the three control groups had any recorded hearing impairments.

3.3 Procedure

We constructed a closed-set identification task. Each of the 48 stimuli was repeated four times, giving 192 tokens in total. These 192 tokens were presented in random order. Subjects heard one token at a time over headphones. They were presented with two answer categories on the computer screen and had to click on the category of the sound they thought they heard. Before the experiment started, the subjects were informed that they would hear Polish sounds.

As answer categories we used the respective Croatian orthographic representations of the sounds. For the sounds which are absent in Croatian, we used the Polish orthographic representations, which fully parallel the Croatian convention, see table 3.

Table 3: Polish and Croatian orthographic representations for the post-alveolar sibilants

	tʂ	**dʐ**	**ʂ**	**ʐ**	**tɕ**	**dʑ**	**ɕ**	**ʑ**
Polish	cz	dż	sz	ż/rz	ć	dź	ś	ź
Croatian	č	dž	š	ž	ć	đ	–	–
used in experiment	č	dž	š	ž	ć	đ	ś	ź

The answer categories for the respective retroflex sounds were always presented on the left side of the computer screen, the answer categories for the respective alveolopalatal sounds always on the right side. For instance, the participant would hear the stimuli [atɕ] and see an answer button labeled "č" on the left side and an answer button labeled "ć" on the right side of the screen.

The Croatian subjects were instructed that Polish has, like Croatian, [tɕ, dʑ, tʂ, dʐ] and additionally "soft" [ɕ, ʑ] in contrast to "hard" [tʂ, dʐ]. The terms "soft" and "hard" are commonly used in Croatian to differentiate between alveolopalatal and retroflex affricates. The sounds were presented to them once before the experiment.

The German and Slovenian subjects heard the sounds twice before the beginning of the experiment. During the experiment they could consult a table with the labels for the answer categories arranged in two rows: retroflex versus alveolopalatal sounds. The German subjects were told that the two sound series they had to distinguish were similar but not identical to two of their native categories, namely the German post-alveolar sibilants [ʃ, ʒ] and the palatal, non-sibilant fricative [ç]. Their answer categories had orthographic representations that were based on these similarities.

Polish subjects were simply asked to categorize the sounds they heard, without a presentation of some of the experimental items prior to the experiment.

In addition to the identification task, the Croatian subjects were asked to read in their every-day variety of Croatian a list of real words. These words displayed the alveolopalatal-retroflex contrast and mirrored the three context and two vowel conditions that we used for the nonsense words in our identification stimuli. The recordings were only checked to provide us with the information whether the participants produced a contrast themselves. For the present paper we did not analyze this data acoustically.

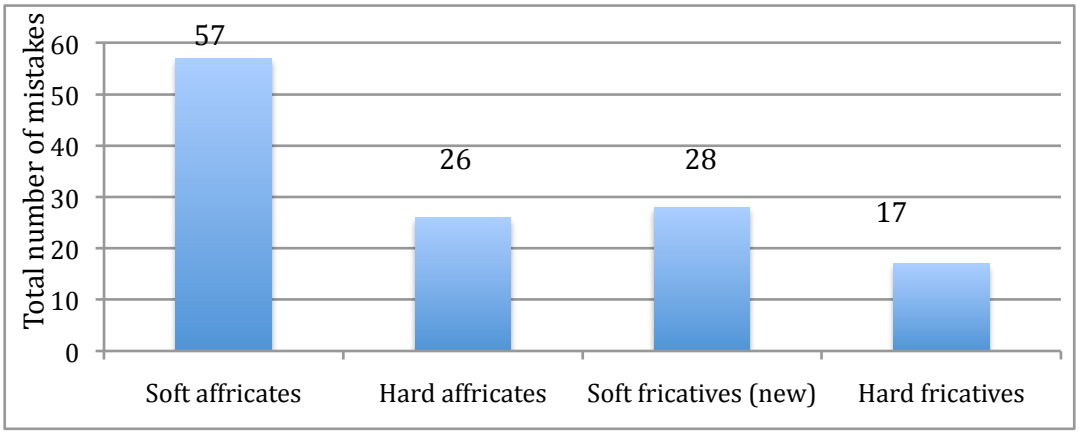
3.4 Results

Mean identification scores for the post-alveolar coronal sibilants were calculated for each group of participants. We discuss the four groups successively, starting with the Croatian participants.

The Croatian native speakers had a mean correct identification score of 96.64%. For the "known" sounds, i.e. the post-alveolar coronal affricates and the retroflex fricatives, they made 100 mistakes, which is an error rate of 3.47%. For the "new/unknown" sounds, i.e. the palatoalveolar fricatives, they made 28 mistakes, which is an

error rate of 2.92%. When split into particular categories, subjects have made 57 mistakes in the identification of alveolopalatal affricates (5.93% of all alveolopalatal affricates), 26 mistakes for retroflex affricates (2.70% of all retroflex affricates), 17 errors in the identification of retroflex fricatives, and 28 mistakes with respect to the new phonemes, i.e. alveolopalatal fricatives (2.91% of all alveolopalatal fricatives). There were no big differences between the two new phonemes: voiceless alveolopalatal was misperceived 16 times, and the voiced alveolopalatal 12 times. There is no significant difference in performance for the known and the new sounds. These results are graphically summarized in (2).

(2) Results of the Croatian subjects

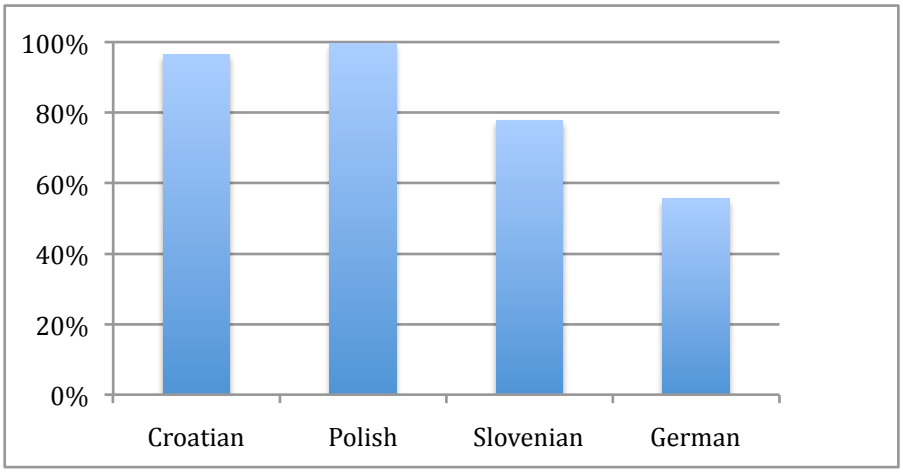


The Polish control group performed at 99.7% accuracy (one subject made one mistake).

The Slovenian control group achieved a mean score of 77.78% correct identification. Within the class of fricatives, they made 17.63% mistakes, and within the class of affricates 27.08%. The error rate for alveolopalatals and retroflexes was almost the same: 10.24% and 11.98%, respectively.

The German control group achieved a mean 55.55% correct identification score. Here we could observe a significant difference between the fricatives and affricates: The participants performed incorrectly for 58.33% of the fricatives, and for 30.55% of the affricates. There are no significant differences between alveolopalatals and retroflex sounds. The comparison of the results of the Croatian subjects and the control groups follows in (3):

(3) Croatian subjects versus Polish, Slovenian and German subjects



4. Discussion

The mean identification scores for the Croatian subjects were in general very high, almost reaching native speaker performance of the Polish control group. We found no significant difference in the performance by the Croatians between Polish sounds that occur as phonemes in their own language and the Polish sounds that are no phonemes in the L1 but are made up of features employed in the L1. These “new” phonemes thus presented

no difficulty in the identification task. Our findings indicate that the phoneme cannot be the ultimate or only phonological unit in perception, otherwise we would have expected the subjects to perform worse for the new phonemes than for the phonemes that occur in their L1. Instead, the present findings seem to provide evidence for a role of phonological features in the perception of sounds: the high identification scores for the new sounds can be explained by the fact that these new sounds are specified by features that the Croatian listeners know from their own phonemes.

The German and Slovenian control groups perform considerably worse than the Croatian subjects. This alone, however, cannot be interpreted conclusively as evidence for the feature or the phoneme as perceptual units. In feature-terms, the German and Slovenian post-alveolar sibilants can be distinguished from their alveolar counterparts simply by the feature [–anterior]. The feature [back] (or any comparable feature) is not employed for the sonorants in these languages (see e.g. Wiese 1996 for German, and Herrity 2000 for Slovenian). The Polish sounds thus could be argued not to be distinguishable by features of their L1, which results in the bad performance. In phoneme-terms, the absence of any of the eight Polish phonemes in both German and Slovenian could be argued to explain the bad performance of both groups.

More support for the feature as perceptual unit comes from the fact that the "new" sounds [ɕ] and [ʑ] occur in Croatian as allophones of the phonemes /ʂ/ and /ʐ/ respectively, that is, the perceptual task requires from the subjects to distinguish between two allophones of the same L1 phoneme. Listeners tend to "overhear" fine perceptual difference between allophones of one phoneme in their native language (e.g., Guion and Lee 2006). In a phoneme-only account, the listeners would therefore be expected not to be able to keep the allophones apart. If we assume, however, that the features are perceptual units independent of the phoneme, then we can explain our findings that Croatians were very well able to keep apart the L2 sounds that are allophones of one phoneme in their L1.

An additional argument for the feature approach might be drawn from the fact that there are no substantial differences in the perception of the two new phonemes, the voiced and voiceless palatoalveolar fricative. If two or more phonemes behave similarly, it is very likely that their common behavior is not just a coincidence but rather results from the fact that these phonemes share a distinctive feature which is an active unit in the perception process.

Within the Croatian group, performances differed considerably. Some subjects performed much worse than the rest, but the maximal error rate was 12%, which is still far below the mean identification error rate of the Slovenian and German control groups (Slovenians 22%, Germans 44%). This highest ratio of mistakes can be observed for Croatian subjects who do not produce a stable contrast between the palatoalveolar and the retroflex affricate series in Croatian themselves: Both subjects with highest error level come from the area of Zadar, where the Čakavian dialect is still spoken (some dialects thereof neutralize the distinction between palatoalveolar and retroflex sibilants). However, the relation between perception and own production is not reciprocal: not all speakers that fail to produce a consistent contrast have a high ratio of mistakes in perception.

The difference in performance between the German and the Slovenian control group is most likely due to the fact that two of the three Slovenians had learned Croatian at school.

5. Conclusions

In our experiment, it has been established that Croats have no problems with the identification of [ɕ] versus [š] and [ʑ] versus [ž], even though they do not have these phonemes in their native language. Compared to the Croats, the German and Slovenian subjects performed far worse on the same task. What Croatian, German and Slovenian subjects have in common is that they do not have alveolopalatal fricatives in their L1 phonemic inventory. They differ, however, in the presence of a phonological feature in their L1 that can potentially differentiate the alveolopalatal and retroflex fricatives: whereas Croatian has such a feature, German and Slovenian does not. We conclude from this that listeners are able to extrapolate their featural knowledge over new phonemes.

We tested the following predictions: If the feature is a unit in L2 phonological perception, "new" phonemes should be perceived as easily as L1 phonemes, as long as the phonological features used to describe the "new" phonemes are used already in the phonological system of the L1. though do not necessary need to mark the distinctions of a given L1 phoneme. This prediction has been borne out. The second prediction said that if a

phoneme is the unit of phonology acquisition, there should be more errors in the perception of "new"/"unknown" phonemes than for known phonemes, and this prediction has not been confirmed. Additionally, the feature approach finds support from two more arguments. First, it is supported by the fact that our subjects could easily discriminate between two L1 allophones by a feature that is used contrastively for other contrasts. Second, we found that the two new phonemes displayed similar rates of misperception, indicating that their similar feature make-up can account for their similar behavior.

The present study could be extended to include measurements of the response time, which might shed more light on the different abstraction levels for speakers with different L1 background German speakers, for instance, are expected to respond much slower in the categorization of the Polish sounds, since they have no abstract phonological categories such as features to help them dealing with this task. A further issue not dealt with here is the perceptual cues for the post-alveolar coronal sibilants in the tested languages. Future research should test the relevance of such cues to see if they can account for the performance differences between, for instance, Slovenian and German participants. Such studies should also include cases where the influence of phonological features and phonetic cues can be strictly separated.

References

Aoyama, Katsura, James Emil Flege, Susan G. Guion, Reiko Akahane-Yamada, Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32: 233-250.

Best, C.T., G.W. McRoberts and N.M. Sithole. 1988. Examination of Perceptual Reorganization for Nonnative Speech Contrasts: Zulu Click Discrimination by English-Speaking Adults and Infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3): 345-360.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Brannen, K. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics*, 47(1/2): 1-46.

Brozović, Dalibor. 1972. O alofonskoj problematici u hrvatskoj ortoepiji (Fonemi i alofoni u standardnome hrvatskosrpskom jeziku). *Radovi. Razdio Lingvističko-Filološki* (6): 5-24. Zadar: Sveučilište u Zadru.

Brozović, Dalibor. 1970. Dijalekatska slika hrvatskosrpskoga jezicnog prostora. *Radovi. Razdio Lingvističko-Filološki* (5): 5-32. Zadar: Sveučilište u Zadru.

Donegan, Patricia. 2001. Constraints and processes in phonological perception. In: Katarzyna Dziubalska-Kołaczyk (ed.), *Constraints and preferences*. Berlin: Mouton de Gruyter. Pages: 43-68.

de Jong, Kenneth and Noah Silbert. 2006. On segmental factorability in second language learning. *IULC Working Papers Online, 6*. http://www.indiana.edu/~iulcwp/

Eckman F., and G. Iverson. 1994. Pronunciation difficulties in ESL: Coda consonants in English interlanguage. M. Yavas, ed., *First and Second Language Phonology*. San Diego, CA.: Singular. Pages: 251-265.

Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15: 47- 65.

Guion, Susan G. and Borim Lee. 2006. The role of phonetic processing in second language acquisition. *English Language and Linguistics*, 21: 123-148.

Hamann, Silke. 2004. Retroflex fricatives in Slavic Languages. *Journal of the International Phonetic Association*, 34:53-67.

Herrity, Peter. 2000. *Slovene: A Comprehensive Grammar*. Routledge.

Lončarić, M. 1996. *Kajkavsko narječje*. Zagreb: Školska knjiga.

Rubach, Jerzy. 1984.*Cyclic and Lexical Phonology. The Structure of Polish*. Dordrecht: Foris.

Silbert, Noah, Kenneth de Jong and Hanyong Park. 2005. Linguistic generalization in L2 consonant identification accuracy: a preliminary report. *IULC Working Papers Online*, 5. http://www.indiana.edu/~iulcwp/

Stampe, David. 1979. *A Dissertation on Natural Phonology*. New York: Garland.

Strange W., Reiko Akahane-Yamada, R. Kubo, S.A. Trent, K. Nishi, and J.J. Jenkins. 1998. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26: 311-344.

Wiese, Richard. 1996. *Phonology of German*. Oxford: Oxford University Press.