Phonological representations and perception of L2 contrasts

Susan Jackson and John Archibald

University of Calgary
sejackso@ucalgary.ca, john.archibald@ucalgary.ca

ABSTRACT

This research investigates factors which underlie the perception of second language (L2) phonological contrasts by highlighting an asymmetry in the perception of the four-way laryngeal stop contrasts in Hindi by native speakers of English and French. We argue that this asymmetry is a result of the influence of differing underlying representations in the first language (L1). Based on a theory that English uses the feature [spread glottis] while French uses the feature [voice] to distinguish voiced and voiceless stops, it was predicted that native speakers of these two language groups would perceive the four-way Hindi contrasts differently.

Monolingual Canadian English (n=18) and monolingual Canadian French (n=18) speakers were tested on their perception of Hindi minimal pairs using an ABX discrimination task with a long interstimulus interval. Results supported the predictions; English speakers performed significantly better on contrasts involving the feature [spread glottis] and the French did significantly better on contrasts involving the feature [voice]. However, perception of all pairs of contrasts by both language groups was above chance, suggesting a role for both phonological interference and phonetic factors.

Keywords: non-native perception, laryngeal contrasts, Hindi

1. INTRODUCTION

1.1 Models of L2 perception

Successful acquisition of phonological contrasts presupposes their accurate perception. While children do this with remarkable ease (e.g. Eimas et al. 1971; Jusczyk 1997), adult learners of a second language are known to have difficulty discriminating between certain sounds that are not employed contrastively in their own language. For example, young children in diverse linguistic environments are able to perceive a contrast between /ɹ/ and /l/, which are contrastive in English but not in Japanese. However, once they begin to construct their native phonological representations, older English children maintain this contrast while Japanese children of the same age are no longer able to perceive it (Goto, 1971; Yamada, 1995).

There is variation in the *degree* of difficulty with which non-native sounds are perceived, which has led to questions about the relationship between the L1 and L2 grammars such as: what factors determine ease of perceptibility, and on what level does the L1 grammar influence the developing L2 grammar? Research in this area has resulted in several models of L2 speech perception. For example, it has been suggested that the degree of difficulty directly relates to the degree of perceived phonetic similarity or dissimilarity between L1 and L2 sounds (see Flege 1995), with more similar phones being the most difficult to perceive, thereby inhibiting the learner from setting up new phonetic categories. Along the same lines, it has been suggested that the perception of non-native contrasts is a process of categorical assimilation (e.g. Best 1995; Best and Tyler 2007) whereby two L2 phones are assimilated to L1 categories based on the perceived similarity between them. Different patterns of assimilation determine the difficulty a non-native listener will have in discriminating the contrast. For example, phones that assimilate to the same L1 category, e.g. dental and retroflex stops for a speaker of English, are the hardest to perceive, while those that assimilate to different categories, e.g. short- and long-lag stops, are easier.

Models in the generative framework argue that the presence or absence of phonological features in the L1 plays a role in the difficulty a learner may have perceiving certain second language speech contrasts (Brown 1998, 2000). This model predicts that if learners lack a particular feature in their L1 grammar that is used to distinguish an L2 contrast, they will be unable to perceive that contrast. However, if the feature in question is represented in learners' L1, regardless of how it is used, they will be able to perceive the L2 contrast employing this feature. For example, Brown (1998) found that despite the lack of phonemically contrastive liquids in Chinese, native Chinese-speaking subjects were able to perceive a contrast between English /l/ and /ɹ/. She argued that this result was due to the presence of the feature [coronal] in their L1, which is the contrastive feature in English liquids. In Chinese, however it is used to contrast fricatives. The native Japanese-speaking subjects, thought to lack this feature in their phonology, were unable to perceive the contrast.

1.2 Representations of laryngeal contrasts

In order to research L1 phonological interference in L2 speech perception, one must integrate it with a theory of phonological feature representation. Determining the correct distinctive feature in an L1 contrast, for example, is crucial to any argument about its influence on the perception of non-native contrasts.

Two types of laryngeal stop systems can be distinguished in the languages of the world: a first type that contrasts prevoiced with short-lag stops, e.g. [b/d/g] vs. [p/t/k], and a second that contrasts short-lag with long-lag stops, e.g. [b̥/d̥/g̊] vs. [p^h/t^h/k^h]. Despite this difference, both these types have traditionally been thought to involve the same distinctive phonological feature, [voice], with the variation in their manifestation being a matter of language-specific phonetic implementation (see Keating 1984). Yet there is compelling evidence that this is not the case, and that the second type (short-lag and long-lag stops) uses the feature [spread glottis] contrastively and not the feature [voice] (e.g. Iverson and Salmons 1995, Kager et al. 2007). Aspiration that accompanies long-lag stops in certain positions is considered to be the phonetic result of a [spread glottis] feature that is part of the phonology. We adopt the position that English is such a language and that it contrasts with French; a true voicing language.

2. STUDY

To investigate the phonological characterization of English and French laryngeal contrasts and their influence on non-native perception, we tested the perception of Hindi laryngeal contrasts by native speakers of these two language groups. While English and French both have a two-way system of laryngeal contrasts, Hindi is a language with a four-way system that has been expressed phonologically as contrasting both [voice] and [spread glottis] (Davis 1995; Avery and Idsardi 2001). Not only does it have voiced, plain and aspirated stops, but it contrasts a fourth type which is both voiced and aspirated, also called 'breathy voice' (Ladefoged and Maddieson 1996), that occurs in neither English nor French.

2.1 Predictions

Assuming a model of L1 feature interference in L2 perception, and taking into account our assumptions about the active laryngeal features of English and French, English speakers were expected to perceive [spread glottis] contrasts well, and have difficulty with [voice] contrasts. Likewise, as French uses [voice] contrastively, French speakers were expected to perceive [voice] contrasts well and have difficulty with [spread glottis] contrasts. Pairs of Hindi stops that contrast by both [voice] and [spread glottis] were expected to be well perceived by participants of both language groups, due to the presence of the relevant laryngeal feature for each language. Predictions are indicated in Figure 1.

Figure 1: Predicted perception of all Hindi contrast types by English and French speakers. ✓ indicates good perception, x indicates poor perception.

	[voice]		[spread glottis]		both features	
Contrast	/b-p/	/b^h-p^h/	/p^h-p/	/b-b^h/	/b^h-p/	/b-p^h/
English	x	x	✓	✓	✓	✓
French	✓	✓	x	x	✓	✓

2.2 Method

An ABX perception task was used to test how well native-English and -French speaking subjects were able to perceive the Hindi voicing and aspiration contrasts. We used a long interstimulus interval of 1500 ms, designed to ensure phonemic judgements, as opposed to acoustic or phonetic (Werker and Logan 1985).

2.3 Participants

The participants were 18 native speakers of Canadian English and 18 native speakers of Canadian French. All were considered to be functional monolinguals, as described by Best and Tyler (2007), in that they were not actively using an L2 or in the process of learning an L2. As functional monolinguals, they were expected to have difficulty categorizing or discriminating phonetic contrasts of non-native languages that are not used to distinguish lexical items in their native language. The English speakers were all residents of Calgary, Alberta while the French speakers were residents of Montreal, Quebec. All participants reported having some experience with either English or French as a foreign language, mainly due to educational policies in Canada that require all children undergo a certain minimum of second language instruction in one of the two official languages. As well, in Montreal, exposure to ambient English through radio, television and in the public sphere is much higher than the same type of exposure to French in many other parts of Canada. In such an environment, it is difficult to establish true monolingualism. However, all participants self-reported as monolinguals in that they had not attained higher than an intermediate level of proficiency in French or English as an L2 in school and reported not to use it in any capacity in day to day life. None of the subjects reported any prior experience with Hindi.

2.4 Stimuli

Natural stimuli were developed with two female native speakers of Hindi who were recruited through the Southern Alberta Heritage Language Association (SAHLA). Hindi words and non-words whose initial stop consonants contrasted by voice, aspiration and place of articulation (dental, alveolar, retroflex and velar) were recorded digitally. All contrastive segments were word-initial and were followed by –/ʌʌl/, e.g. /pʌʌl, p^hʌʌl, bʌʌl, b^hʌʌl/, in order to create minimal pairs.

Each speaker recorded words singularly as well as in a carrier sentence, which was presented in Devanagari script. Stimuli for the experiment were chosen based on similar intonation, to avoid non-relevant variation influencing the participants' judgement of similarity. A total of 192 trials, counterbalanced in terms of order of words, targets and speakers, were randomized and presented in six blocks of approximately five minutes each.

2.5 Task

For each trial, participants listened to a series of three words and were told that the first two words (the minimal pair AB) differed by the first consonant, to avoid any confusion about other possible phonetic differences, such as variation in the length of the vowel. Participants were asked to judge if the third word (X) in the series was most like the first word they heard or the second. As an example, they were presented with a sequence such as (A) *paal*, (B) *baal*, and (X) *paal*. If they judged (X) to be most like the first (A) or most like the second (B), they were instructed to press number 1 or 2 respectively. A result of correct or incorrect and the speed at which they responded were displayed on screen after each trial.

3. RESULTS

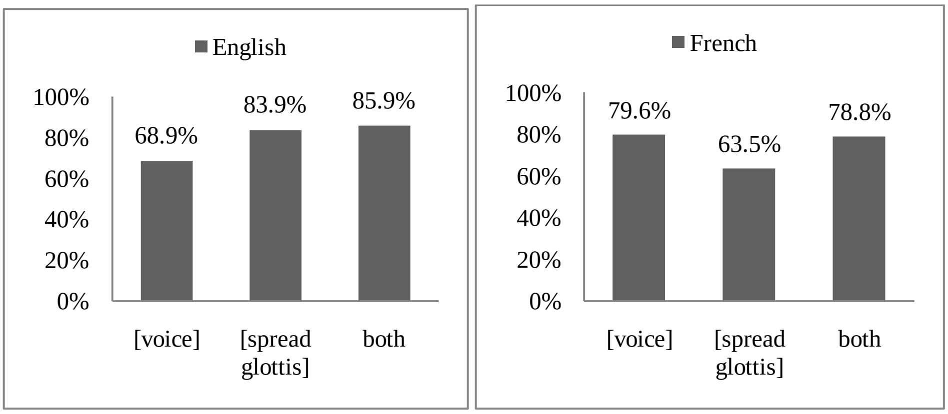
Data was analyzed using a multivariate analysis of variance (ANOVA) of the percentage of correct responses by language group. Comparisons were made between and within groups. The confidence interval was 95%.

On all pairs of stops that contrasted by [voice] only, the French subjects performed significantly better than the English subjects [$F(1, 34) = 20.05$, $p<0.05$], while on all pairs that contrasted by [spread glottis] only, the English did significantly better than the French [$F(1, 34) = 64.15$, $p< 0.05$].

The graphs in figure 1 below show that the English subjects performed similarly on pairs that contrasted by [spread glottis] and those that contrasted by both [spread glottis] and [voice], with 83.9% and 85.9%

correct responses respectively. The difference was not statistically significant (p=0.156). Therefore, it would appear their performance was unaffected by cues to contrastive [voice] in the Hindi stimuli. Likewise, regardless of the presence of aspiration, the French subjects performed quite similarly on the [voice] contrasts and those pairs that contrasted by both features with 79.6% and 78.8% correct responses respectively. Again, the difference was not statistically significant (p=0.768).

Figure 1: Average discrimination scores for pairs grouped by contrastive feature.

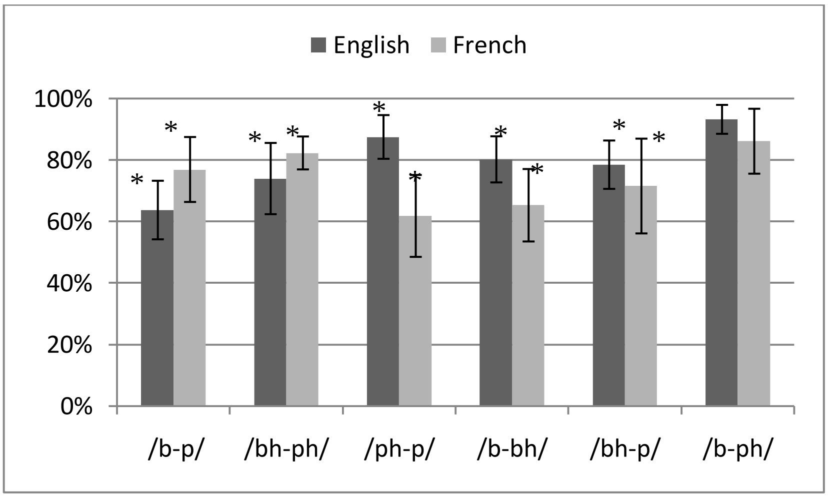


We predicted the same pattern to hold for individual pairs contrasting by these features, such that the French speakers were expected to distinguish the /b-p/ and the /b^h-p^h/ pairs more easily than the English speakers, as these pairs contrast by the feature [voice]. It should be kept in mind that /b/ refers to the Hindi prevoiced /b/. For /p^h-p/ and /b-b^h/ pairs, the opposite was predicted: the English speakers were expected to distinguish them better than the French, as these pairs contrast by the feature [spread glottis]. Performance on pairs that contrast by both features was expected to be good for both participant groups. The results from the experiment can be seen in Table 2, followed by a graph in Figure 2.

Table 2: English and French average discrimination scores for individual pairs. Numbers represent percentage of correct responses with standard deviations in parentheses. /b/= voiced, /b^h/= voiced, aspirated, /p/= voiceless, unaspirated (or plain), and /p^h/= voiceless, aspirated

	[voice]		[spread glottis]		both features	
Contrast	/b-p/	/b^h-p^h/	/p^h-p/	/b-b^h/	/b^h-p/	/b-p^h/
English	63.7 (9.6)	74.0 (11.6)	87.5 (7.1)	80.2 (7.5)	78.5 (7.9)	93.2 (4.7)
French	76.9 (10.6)	82.3 (5.4)	61.8 (13.3)	65.3 (11.8)	71.5 (15.4)	86.1 (10.6)

Figure 2: English and French average discrimination scores for individual pairs. Error bars represent standard deviation.



For all pairs that contrast by a single feature, there was a significant difference in performance between the subject groups, confirming the main predictions. However, looking at those pairs that contrast by both features, there was a significant difference in performance between groups on results for the /b-p^h/ pair, while there was no significant difference for the /b^h-p/ pair, [$F(1, 34) = 2.90$, $p<0.05$].

4. DISCUSSION AND CONCLUSION

The present study investigated the effect of differing underlying laryngeal representations on English and French speakers' perception of the Hindi four-way laryngeal contrasts. Participants' native language had a significant effect on their ability to distinguish between these contrasts. A clear asymmetric pattern of perception between language groups emerged: English speakers were able to perceive contrasts that employ the [spread glottis] feature more readily than contrasts that employ only [voice], while the reverse was found for French speakers.

However, upon examination of the results for individual contrastive pairs, additional complexities were found. For example, there was no significant difference between the results for the English speakers on the pairs /b^h-p^h/ (74.0%) and /b^h-p/ (78.5%), (p = 0.12), yet the former contrasts the [voice] feature only. Also, there was no significant difference between the results for the French speakers on the pairs /b-b^h/ (65.3%) and /b^h-p/ (71.5%), (p=.08), yet the former contrasts only the [spread glottis] feature. What is notable here is that deviations from the predicted pattern centre on voiced aspirated stops; the segment that is novel to both languages yet employs distinctive features from each.

It is possible that the presence of certain cues in the voiced aspirated stops may aid English listeners to better distinguish between a voiced aspirated and a voiceless aspirated stop despite the lack of a [spread glottis] contrast. Archibald (2005) proposes that second language learners are able to acquire phonological features that are absent from their L1 when the acoustic cues are perceptually robust. As such, their strong auditory saliency and resistance to masking (Wright, 2004) are able to override L1 phonological filters.

The individual French contrasts conformed to predictions with the exception of /b^h-p/. Interestingly, the presence of aspiration appeared to lower the ability to perceive this contrast (71.5%) for this group despite the different in voicing, although it is possible poor discrimination was based on other factors.

Finally, while Brown's Feature model of perception predicts that the absence of a feature in the L1 should prohibit discrimination of certain contrasts, all subjects performed above chance on all contrasts. This suggests that a strong interpretation of this model cannot be supported. A more likely interpretation is one in which multiple factors – phonetic and phonological – are involved in the discrimination of non-native contrasts.

5. REFERENCES

Archibald, J. 2005. Second language phonology as redeployment of L1 phonological knowledge. *Canadian Journal of Linguistics* 50: 285 -314.

Avery, P. and Idsardi, W. J. 2001. Laryngeal dimensions completion and enhancement. In: Hall T. (ed.), *Distinctive feature theory*. Berlin: Mouton de Gruyter. 41-70.

Best, C. T., and Tyler, M. D. 2007. Non-native and second-language speech perception: Commonalities and complementarities. In: O.-S. Bohn and M. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. 12-34. Amsterdam: John Benjamins.

Best, C.. T. 1995. A direct realist view of cross-language speech perception In: Strange, W. (ed.), *Speech perception and linguistic experience: issues in cross-language research*. York: Timonium. 171-204.

Brown, C. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In: John Archibald (ed.), *Second language acquisition and linguistic theory*. Oxford: Blackwell. 4-63.

Brown, C. 1998. The role of the L1 grammar in the acquisition of L2 segmental structure. *Second Language Research* 14: 136-193.

Davis, K.. 1995. Phonetic and phonological contrasts in the acquisition of voicing: Voice Onset Time production in Hindi and English. *Journal of Child Language* 22: 275-305.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., Vigorito, J. 1971. Speech perception in infants. *Science.* 171: 303–306.

Flege, J. 1995. Second language speech learning: Theory, findings, and problems. In: Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press. 233-73.

Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9: 317-23.

Iverson, G. and Salmons, J. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12: 369–96.

Jusczyk, P. W. 1997. *The discovery of spoken language*. Cambridge, MA: MIT Press.

Kager, R., van der Feest, S., Fikkert, P., Kerkhoff, A., Zamuner, T. 2007. Representations of [voice]: Evidence from acquisition. In: van de Wiejer, J. M. And van der Torre, E. J. (eds.), *Voicing in Dutch*. Amsterdam: John Benjamins, 41-80.

Werker, J. and Logan, J. 1985. Cross-Language Evidence for Three Factors in Speech Perception. *Perception and Psychophysics*. 37: 35-44.

Wright, R. 2004. A review of perceptual cues and cue robustness. In: Hayes, B., Kirchner, R., Steriade, D. (eds.), *The phonetic bases of markedness*. Cambridge: Cambridge University Press. 34-57.

Yamada, R. A. 1995. Age and acquisition of second language speech sounds: perception of American English /g/ and /l/ by native speakers of Japanese. In: Strange, W. (ed.), *Speech perception and linguistic experience: issues in cross-language research*. York: Timonium. 305-320.