# CATEGORIZATION OF BILABIAL STOP CONSONANTS BY BILINGUAL SPEAKERS OF ENGLISH AND POLISH

Elzbieta B. Slawinski*, Megan Wiigs*
*Department of Psychology, University of Calgary, Canada

## ABSTRACT

The range of VOT characterizing voiced [b] and voiceless [p] stop consonants is different in English and Polish languages. The study focused on the influence of voicing distinction in Polish on that in English.

The experiment included 10 monolingual English listeners and 10 bilingual Polish listeners who had learned to speak English in an intensive setting between six and twelve years of age. Participants listened to a [bi-pi] continuum of 10 synthesized syllables with VOT values varying from -40 to +50 ms. Ten repetitions of each stimulus were presented to the listeners within Polish and English settings.

Results indicated that the location of the phonemic boundary between categories of voiced and voiceless stop consonants differed significantly between bilingual and monolingual groups. Moreover, a significant influence of the language setting on categorization of stop consonants by bilinguals suggests that they may possess a distinct phonetic inventory representation for each language.

## 1. INTRODUCTION

In many languages Voice Onset Time (VOT) is a variable that acts as an acoustical cue in the differentiation between the bilabial stop consonants: voiced [b] and voiceless [p]. It is known that the range of VOT that characterizes voiced and voiceless stop consonants is different in English and Polish languages. In Polish, the VOT for the phonemic boundary between the sounds [b] and [p] (in a word-initial position) is located slightly above 0 ms. However, for English speakers, a stop consonant on the right side of the Polish boundary is still perceived as [b] until the higher VOT value of 30 to 35 ms [8].

With respect to phonological representation in bilingual speakers, previous research has led to the development of two major hypotheses. The first is that bilingual speakers may adopt intermediate production and perceptual values as a type of compromise between the phonemic boundaries of their first language and second language [7, 11].

Other studies have proposed that bilingual individuals can develop two separate phonetic systems [4, 5]. A limited number of speech perception studies have manipulated language "sets" to determine if bilingual speakers establish distinct phonetic representations for the sounds of their second language. Distinct language contexts have been created, for example, by changing the language in which instructions or extraneous conversations are presented. It is hypothesized that if bilingual listeners possessed distinct representations for the sounds of their first and second languages, they would show a "switching" of phonemic boundaries, depending on the language set.

To date, the results of these studies have been relatively inconclusive. Many studies have failed to show significant perceptual effects according to perceptual set [2, 3, 10]. Significant differences according to perceptual set were found by Flege and Eefting (1987a), who recorded changes in the perceptual boundaries of Dutch-English bilinguals. These researchers used more rigorous procedures in attempting to induce participants to process stimuli as though they were listening to two different languages. Overall differences were significant, but were also much smaller than would be expected if one were to compare monolingual speakers of both languages. Bohn and Flege (1993) also found a significant language set effect. However, similar effects were noted in a monolingual control group, thus rendering their results questionable at best.

The goal of this study was to explore the influence of a first language (Polish) on the acquisition of a second language (English). In particular, the study focused on the influence of voicing distinction in the first language on the categorization and underlying phonological representations of bilabial stop consonants in the second language. Thus, the main focus of the present research was to examine differences in the categorical perceptions and phoneme boundary locations of bilingual speakers of both Polish and English, when exposed to stimuli presented in two different language sets. Results are compared to those of monolingual speakers of English.

## 2. METHOD AND PROCEDURE

### 2.1. Participants

Two groups participated in this study. The control group (N = 10) was composed of monolingual English listeners, the experimental group (N = 10) of bilingual listeners. Members of the bilingual group were native Polish speakers who had learned to speak English in an intensive setting between six and twelve years of age. These participants had similar types and amounts of the second language exposure in that they were all born in Poland and moved to Canada. All bilingual participants continued to use Polish on a regular basis at home and in some social settings, while using English as a main language at University or secondary school.

### 2.2. Stimuli

Stimuli were synthesized using KLSYN 88a parallel/cascade software. Participants listened to a continuum [bi-pi] of synthesized syllables varying from -40 to +50 ms VOT values in steps of 10 ms each. All other parameters (e.g., burst amplitude, formant transitions) remained constant. Each syllable was approximately 500 ms long. This continuum covered the range from the prevoiced, negative

VOT bilabial stop consonant [b], to the voiceless, long-lag [p]. Ten randomly-ordered presentations of each of the ten different stimuli (syllables) were presented in the contexts of perceptual sets in each of Polish and of English. The different perceptual sets were created by presenting pre-recorded instructions and questions read by native Polish and English speakers. Within each set, participants heard between 1.5 and 2 minutes of speech. Participants were also asked to read loudly 8 sentences written either in Polish or in English. These sentences included words with [b] and [p] in the initial position.

### 2.3. Procedure

For each stimulus presentation, participants were required to identify syllables as [bi] or [pi], using an alternative response box labeled with large upper case letters B and P. Moreover, participants recorded their subjective feelings about the quality or intelligibility of the presented syllable on a three-choice rating scale. Participants rated the quality of the syllable by marking one of three possibilities (happy face - good quality, neutral face - neutral quality, and sad face - poor quality). All instructions, questions, and [bi] and [pi] stimuli were presented binaurally over headphones at a sound pressure level of 75 dB.

## 3. RESULTS

Percentages of either [bi] or [pi] responses were computed for each participant and each language set. Figure 1 presents a summary of the identification response distributions according to group and perceptual set.

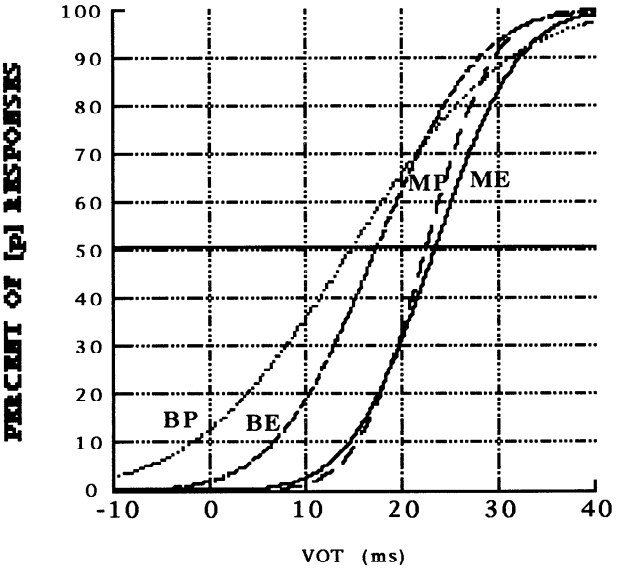


Figure 1. Response identifications as a function of VOT for bilingual (B) and monolingual (M) listeners obtained in the Polish (P) and English (E) perceptual sets. Perceptual crossover from [b] to [p] category (50% [p] responses) for bilinguals listeners in the Polish (BP) and English (BE) sets is associated with different VOT in the Polish (BP) and English (BE) sets. The crossover is located at the same VOT for monolingual English listeners (MP and ME).

| Subject's ID | Sex | Age (Years) | Age Learned English (Years) | Years in Canada | Phonemic Boundary (ms) (Polish set) | Phonemic Boundary (ms) (English set) |
|---|---|---|---|---|---|---|
| B1 | F | 21 | 6.5 | 14.5 | 13.85 | 20 |
| B2 | M | 21 | 10 | 11 | 11.7 | 11.7 |
| B3 | F | 24 | 8 | 16 | 7.3 | 8.14 |
| B4 | M | 23 | 7.5 | 15.5 | 18.2 | 13.5 |
| B5 | M | 23 | 8 | 15 | 22.99 | 28.85 |
| B6 | M | 30 | 12 | 16 | 17.88 | 16.96 |
| B7 | F | 21 | 6 | 15 | 18.67 | 19.59 |
| B8 | M | 13 | 9 | 4 | 11.54 | 25 |
| B9 | F | 20 | 12 | 8 | 12.12 | 15 |
| B10 | F | 14 | 7 | 7 | 14.37 | 15.92 |
| Mean | | 21 | 8.6 | 12.2 | 14.86 | 17.47 |

Table 1. Mean phonemic boundaries for two perceptual sets of bilingual listeners

Two category boundary scores were computed for each participant. These corresponded to identification responses in the Polish and English perceptual sets, and mean values are summarized for bilingual participants in Table 1.

The category boundary scores indicate the location on the VOT continuum where half of the stimuli were identified as the beginning of a syllable with [b] and half with [p]. In other words, this is where the perceptual crossover between voiced and voiceless sounds occurred for each participant.

There were no significant differences in the category boundary scores of monolingual speakers in the two language environments. A significant difference ($t(9)=2.098$, $p<0.05$) was observed in the category boundaries of bilinguals listening to English ($M=17.47$, $SD=6.17$) versus Polish ($M=14.86$, $SD=4.57$).

Between-subject tests compared the results of the monolingual and bilingual groups in each perceptual sets. First, there was a significant difference ($t(18)=2.467$, $p<0.05$) between the category boundary scores of the monolingual group in the English perceptual set ($M=22.63$, $SD=4.6$) versus the bilingual group in the same set. Next, there was a significant difference ($t(18)=4.109$, $p<0.05$) in the boundary values of the monolinguals listening to Polish ($M=22.26$, $SD=4.13$) versus the bilinguals listening to Polish.

Although to date we have not applied detailed statistical analyses to the quality rating of syllables' data, the syllables at either extreme of the VOT continuum were rated by both groups of listeners as more intelligible (good) than the mid-range stimuli. The rating of the mid-range stimuli differed between the bilingual and monolingual listeners. In particular the stimulus, which corresponded to 10 ms VOT was identified by both groups of listeners in both perceptual sets as [bi], but was rated by majority of bilinguals as being of poor quality and by majority of monolinguals as being of good quality (see Figure 2).

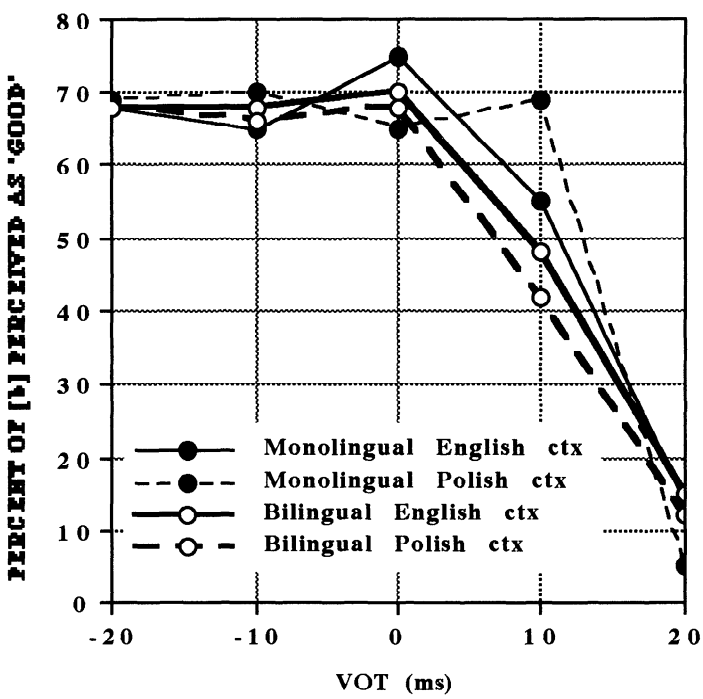


Figure 2. Mean percentage of [bi] perceived as "good" by bilingual and monolingual listeners in the contexts (ctx) of Polish and English perceptual sets.

## 4. DISCUSSION

For the present study it was hypothesized that the two different testing situations created by the Polish and English perceptual sets would alter the responses of bilingual speakers, but not those of monolingual English speakers.

Analyses of the data showed that this hypothesis was supported. Monolingual speakers showed no significant differences in category boundary scores in a comparison of results from each perceptual set, and their phonemic boundary was found to be within a range of "typical" VOT

for English listeners. This is an important finding in that it serves to strengthen the data obtained from bilingual speakers. Because monolingual English speakers showed no change in the location of phonemic boundary according to perceptual set, we can conclude that simply listening to speech streams containing different VOT values was not sufficient to cause perceptual boundary changes, as may have been the case in the study conducted by Bohn and Flege (1993).

In the present study, bilingual listeners showed significant differences in their categorization of voiced and voiceless stop consonants in the two language environments. The presence of this small, yet significant shift in the bilinguals but not in the monolinguals suggests that the manipulation of perceptual sets was successful in systematically altering the boundary locations for bilingual speakers. Similar to Bohn and Flege (1993), we can infer that the bilinguals had entered into different processing modes in the two language sets from the fact that the language set questions were answered sensibly. Bilingual speakers thus perceive acoustically identical stimuli differently, depending on the language they are processing at that time. This perceptual shift in phonemic boundary location according to the language environment suggests that bilingual listeners may possess a distinct phonetic inventory representation for each language.

The present study is one of very few that compare speech perception and categorization in monolingual speakers of English and bilingual speakers of both Polish and English. Of greatest interest in this experiment is the finding that differences in identification occur not only between speakers of different languages (i.e., the VOT values of the phonemic boundaries between categories of voiced and voiceless stop consonants differed significantly between bilingual and monolingual groups) but also within multilingual speakers. Thus it appears that either separate phonetic categories for [b] and [p] in each language are established by bilingual speakers in each language, or else that these speakers may adopt very flexible intermediate values.

In a previous study, Bohn and Flege (1993) noted that although stimuli were identified as belonging to either [t] or [d] categories, these may have sounded "distorted or accented". The present study incorporated a three choice rating scale in order to examine the reactions of two groups of participants to the quality of syllables of synthesized continuum. It has been found that the bilingual listeners, independently of the perceptual set, were qualifying the [b] consonant of longer VOT as poor in quality. On the other hand the [b] consonant corresponding to the same VOT was perceived as being of good quality by the monolingual listeners of English. Thus it might be that bilinguals and monolinguals are not using the same phonetic inventories.

However, due to limitations in the design of the present study, we cannot explain why bilinguals are shifting the phonemic boundaries between voiced and voiceless consonants depending on the language environment. Do they process phonemes in two separate phonemic inventories or do they search for an appropriate template within only one phonemic inventory, with the template satisfying the fuzzy logical model of speech perception [9]? Future studies should explore in depth the phonetic inventory representation of bilinguals.

As a group, bilinguals showed evidence of a significant perceptual shift. Individually, however, this effect was not always seen. Future studies will help to determine what factors are involved in this phenomenon of perceptual switching. For example, it is likely that factors including age when the second language was acquired, and quantity and quality of second language exposure, affect ways in which syllables are categorized by bilingual speakers. A follow-up study currently underway at the University of Calgary will attempt to address potential interactions of speech perception with age of second language learning. It is our expectation that further knowledge about how language is represented in the cognitive network of the bilingual speaker will help to understand both bilingualism as well as language representation, processing, and retrieval in general.

## References


[1] Bohn, O., & Flege, J. E. 1993. Perceptual switching in Spanish/English bilinguals. *Journal of Phonetics*, 21, 267-290.

[2] Caramazza, A., Yeni-Komshian, G., & Zurif, E. B. 1974. Bilingual switching: The phonological level. *Canadian Journal of Psychology*, 28, 310-318.

[3] Elman, J. L., Diehl, R. L., & Buchwald, S. E. 1977. Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, 62, 971-974.

[4] Flege, J. E. 1988. The production and perception of speech sounds in a foreign language. In Winitz, H. (ed.), *Human communication and Its disorders*, Norwood, NJ: Ablex.

[5] Flege, J. E. 1990. The intelligibility of English vowels spoken by British and Dutch talkers. In R. Kent (ed.), *Intelligibility in speech disorders: Theory, measurement, and management.* Amsterdam: Benjamins.

[6] Flege, J. E., & Eefting, W. 1987a. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185-202.

[7] Flege, J. E. & Eefting, W. 1987b . Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, 67-83.

[8] Keating, P.A., Mikos, M.J., & Ganong, W.F. II 1981. A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America*, 70, 1261-1271.

[9] Massaro, D.W. 1989. Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, 21, 398-421.

[10] Williams, L. 1977. The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics*, 21, 289-297.

[11] Williams, L. 1979. The modification of speech perception and production in second-language learning. *Perception and Psychophysics*, 26, 95-104.