PERCEPTION OF PREVOICED STOP CONSONANTS BY MONOLINGUALS AND BILINGUALS: EVIDENCE FOR DIFFERENT PERCEPTUAL SENSITIVITIES

L. Coixao and N. Bacri

Laboratoire de Psychologie Expérimentale
Paris V-CNRS-EHESS,Paris, France.

Listeners assimilate foreign speech sounds to their own phonemic categories whenever possible. But what happens for bilinguals when their two languages are closely related? French monolinguals (MF) and Portuguese-French bilinguals (BPF) were tested in identification and AXB discrimination tasks. MFs'fast responses were non random except for the longest prevoicing, whereas BPFs' showed two peaks around Portuguese and French referential values. According to acoustic patterns and task demands, listeners rely either on a phonemic processing strategy or on a goodness of fit strategy which allows MF to build an allophonic space and BPF to keep separate their two languages.

While there is evidence that linguistic experience affects the ability to process phonemic categories as early as the last quarter of the first year [7], there is some disagreement about whether the perceptual analysis bilinguals have to perform is thoroughly determined by phonological constraints or not. When the two languages are closely related, they are not differentiated at the phonological level [2]. But under certain conditions, the effect of phonological constraints can be weakened, and listeners can rely on phonetic cues to keep separate their perceptual representations [3].

Allophonic variants from the point of view of phonemic labelling can be perceived as different. It has been hypothesized that discrepancies between native and nonnative speech sounds receiving an identical label are processed with reference to the acoustic distance between any exemplar and the category center [4] i.e. the acoustic configuration usually produced by native speakers of the two languages.

The present experiments study how listeners process perceptual dissimilarities in two cross-language situations: perception of a Portuguese /da/-/ta/ VOT continuum by French monolinguals (MF) and by Portuguese-French bilinguals (BPF). VOT is generally considered as the most salient cue for voicing when opposing voiced and unvoiced categories if not prevoiced and voiced ones [5]. As for the stop consonant subset, French and Portuguese are closely related. Both languages present a prevoiced-voiced contrast, opposing a long (French) or a very long (Portuguese) voicing lead to a null or a short lagging VOT. According to the assimilation hypothesis [1], allophonic processing for foreign, but neighbouring sounds such as those we study here, is phonemic. MF will assimilate all the prevoiced stimuli to the /da/ category. By contrast, if category goodness plays a role, it could limit allophonic space to certain stimuli.

But what happens for bilinguals when their two languages are closely related? Are their two languages differenciated at the level of perceptual representations, allowing them to detect phonetic differences related to their two languages within a single phonemic category [3]? In this case, they should exhibit a good discrimination accuracy for two distinct areas, corresponding either to the /da/-/ta/ boundary or to a contrast between the French and Portuguese /da/ phonetic categories. On the contrary, if they assimilate the members of the voicing contrast in one of the two languages to those of the other language, due to a partial acoustic overlap, their discrimination should be random, except in the /da/-/ta/ boundary area, common to both languages.

1. EXPERIMENT 1

This experiment was designed first to determine the phonemic /da/-/ta/ boundary values, second to study whether a shift, marking interlanguage interferences [2,3], appeared between BPF and MF responses or not.

1. 1. Method

Subjects. The subjects were 5 MF and 5 BPF students with normal hearing. BPF first language was Portuguese. All bilinguals had been living in France since at least 15 years and acquired French before the age of 5.

Stimuli. A /da/ syllable, produced by a Portuguese monolingual female, was selected (syllable duration: 276 ms, VOT: -96 ms). The test stimuli were digitized at a 16 KHz sampling frequency and VOT reduced by 12 ms steps (from -96 ms to -36 ms) or 6 ms steps (from -36 ms to 0 ms) along the /da/-/ta/ VOT continuum.

Procedure. Subjects listened individually over earphones, in a quiet room, at a comfortable listening level, to 10 blocks, each consisting of one complete randomization of the continuum. The ISI was 3 s and the IBI was 20 s. Listeners' responses were forced choice "Da" or "Ta". All instructions were given in French.

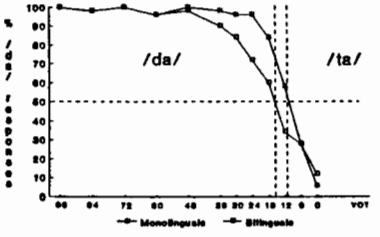

Figure 1. Identification functions for MF (•) and BPF (○) listeners on the /da/-/ta/ continuum (in ms).

1. 2. Results and discussion

The average labeling functions for the two groups are plotted in Figure 1. The /da/-/ta/ boundary fell at -9.2 ms of prevoicing for MF and at -16.6 ms for BPF. An Anova on boundary values showed that this difference between groups was significant ($F(1,8) = 10.4$, $p < .02$). The steep curves suggest that a leading VOT is a strong perceptual cue for BPF as well as for MF. These results differ from those of previous studies on prevoicing [2, 5]. Moreover, there is a clear shift between MF and BPF identification functions. Whatever the case, identification data support the hypothesis of an assimilation of allophonic phonemic variants.

2. EXPERIMENT 2

Even though Experiment 1 suggested an assimilatory process, forced choice labeling could have interfered with perception of differences between stimuli. If allophonic variants have been perceptually assimilated, both MF and BPF should have a good discrimination accuracy just for the stimuli spanning their respective phonemic boundary. Should MF data be non random on the long lead end of VOT continuum and BPF around the medium VOT values, it would undermine assimilation hypothesis and suggest a multi-level processing.

2. 1. Method

Subjects. 10 MF and 10 BPF were tested.

Stimuli and procedure. The same 12 stimuli as in Experiment 1 were used in an AXB discrimination task. A training block of 32 trials preceded 5 blocks of 36 trials, randomized within blocks. ISI was 500 ms, ITI 4 s and IBI 20 s. Subjects had to respond, as quickly and accurately as possible, whether the X stimulus was the same as the first or the third stimulus, by pressing one of two buttons.

2. 2. Results

Mean values of correct responses for the two groups are plotted on Fig. 2. Each data point corresponds to 200 responses per group. The discrimination function for MF exhibited a maximum on the rightward end of the continuum, suggesting an effect of the phonemic /da/-/ta/ boundary. But correct responses are clearly above chance from pair 5 onwards (binomial test, $p < .001$). Results for BPF were less clear-cut, as their discrimination function showed just a

slight peak around a 20ms prevoicing value.

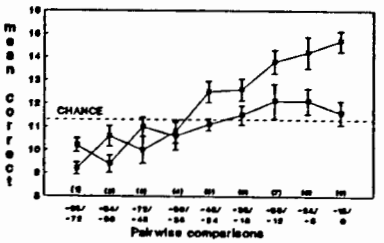

Figure 2. Correct discriminations for MF (•) and BPF (o) as a function of stimulus pairs.

An ANOVA on the correct responses showed an effect of stimulus pair just for MF ($F_{(8,72)}$= 6.9, p<.0001). BPF responses were significantly more correct for the four stimulus pairs presenting at least one short prevoicing (pairs 6 to 9) than for the other ones ($F_{(1,9)}$=8.55, p<.01). Between-group difference was significant for the rightward end of the continuum ($F_{(1,18)}$=6.95, p<.01).

What suggests first a link between discrimination accuracy and phonemic boundary: Discrimination is all the more correct as stimuli pairs span the phonemic boundary. Second, the difference between MF and BPF for the pair enclosing the null VOT value confirms that the slight leftward shift of BPF responses is significant.

An ANOVA on RT data showed a main effect of stimulus pair (MF: $F_{(8,72)}$=6.1, p<.0001; BPF: $F_{(8,72)}$=2.4, p<.03), but not of subject group. Mean RT for MF was 733ms (sd=144ms), for BPF 772ms (sd=202ms).

Following [6], we carried out a threefold partition of RT data to specify the time course of discrimination processes. Each subset/subject contained a third of the data. Between groups range differences were nonsignificant. Proportions of correct responses (Fig. 3a and 3b) were computed for each pair and averaged across subjects (% correct for a specific RT partition * % correct for each partition relatively to the set of correct responses).

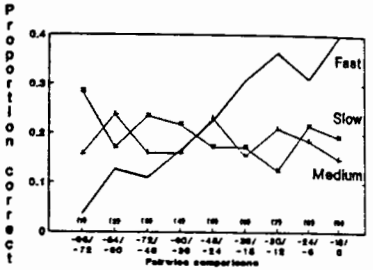

Figure 3a. Proportion of correct discriminations for each RT partition. Monolingual data.

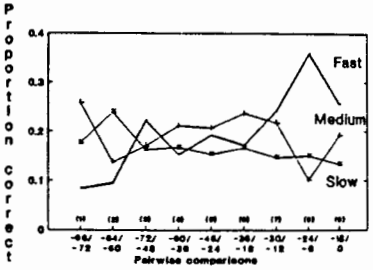

Figure 3b. Proportion of correct discriminations for each RT partition. Bilingual data.

An ANOVA on these proportions showed a main effect of stimulus pair ($F_{8,144}$=4.4, p<.0001). The main difference between groups concerned fast RT (p<.02). Whereas MF discrimination function was quasi-linear, BPF data exhibited 2 peaks, on the pairs 3 and 8. Both groups discriminated better the pair straddling their respective phonemic boundary, but in addition MF showed a good accuracy for "intracategory" pairs and BPF for a pair opposing a long prevoicing to a medium-sized one. It is worth noting that medium and slow responses were random for BPF, but above chance near the rightward end of the continuum (medium and slow RT) and the leftward end (slow RT) for MF.

3. DISCUSSION

A comparison between identification and discrimination data indicates that when processing either a neighbouring language or their two languages, listeners use the phonemic contrast between a long and a short or null prevoicing, but not in all conditions. Specifically, either when labeling stimuli or when responding as fast as possible, they rely on VOT, the critical value of which is common to both languages. Bilinguals and monolinguals gave the same pattern of responses, with a bilingual boundary shift leftward.

However, data do not permit to conclude that both groups have assimilated Portuguese to French stop consonants, as predicted by the assimilation hypothesis [1]. First, MF showed a discrimination accuracy that exceeded widely the /da/-/ta/ boundary area, specially when fast. Second, bilinguals' fast responses were above chance for the contrast between a very long and a medium prevoicing. Thus, listeners of both groups have detected phonetic differences between phonemes receiving the same phonemic label, when corresponding either to values usually produced (MF) or to the contrast between values that respectively characterize bilinguals' two languages. The "goodness of fit" of one of the stimuli in the pair may have facilitated accuracy, if it has been used as a referential point in consideration of French or of French and Portuguese languages.

Another striking result is that, in an AXB test, BPFs' slow responses were never above chance, even though their discrimination accuracy is not significantly poorer than MFs'. In a task having high memory requirement listeners may rely not only on the more salient cue, but also on all the potential cues [8]. Assuming that in this case perceptual analysis takes more time to be processed, its issue depends mainly on the compatibility of cues. Should multiple cues be perceived as lacking coherence, their analysis could not result in a strong discrimination accuracy. It is what happens to bilinguals who, speaking equally well both languages, are plausibly sensitive to the discrepancy between temporal and spectral cues e.g. shorter and shorter French-like prevoicing vs. Portuguese formant values.

Our data thus provide some support to the hypothesis that listeners can use distinct processing strategies when identifying and discriminating speech syllables. According to the acoustic patterns and task demands, they rely on a phonemic processing strategy, specifically in the phonemic boundary area. However, they may take into account the category goodness, when farther from this boundary, in order to differentiate syllables receiving the same label. They can build an allophonic space, and bilinguals can keep separate their two languages. Thus bilinguals and monolinguals appear as perceiving speech according to the same processes, but with different perceptual sensitivities due to linguistic experience.

REFERENCES

[1] BEST, C. (in press), "The emergence of language-specific phonemic influences in infant speech perception", in C. Nusbaum and J. Goodman (eds.), *Development of speech perception*, Cambridge: MIT Press.

[2] CARAMAZZA A., YENI-KOMSHIAN G., ZURIF E., CARBONE M.E. (1973), "The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals", *J.A.S.A*, *54*, n°2, 421-428.

[3] FLEGE, J. (1991), "Age of learning affects the authenticity of VOT in stop consonants produced in a second language", *J.A.S.A.*, *89*, 395-411.

[4] GRIESER, D., KUHL, P. (1989), "Categorization of speech by infants: Support for speech-sound prototypes", *Develop. Psychology, 25*, 577-588.

[5] KEATING P., MIKOS M., GANONG III W. (1981), "A cross-language study of VOT in the perception of initial stop voicing", *J.A.S.A*, *70*, 1261-1271.

[6] MILLER J., DEXTER E. (1988), "Effects of speaking rate and lexical status on phonetic perception", *J. of Exp. Psychology: H.P.P, 14*, 369-378.

[7] WERKER J., TEES R. (1984), "Phonemic and phonetic factors in adult cross-language speech perception", *J.A.S.A*, *75*, 1866-1878.

[8] WHALEN, D., SAMUEL, A. (1985), "Phonetic information is integrated across intervening nonlinguistic sounds", *P. & P., 37*, 579-587.