

A limit on behavioral plasticity in speech perception

Christophe Pallier

Departament de Psicologia Bàsica, Universitat de
Barcelona, Barcelona, Spain and Max-Planck-Institute for
Psycholinguistics, Nijmegen, The Netherlands

Laura Bosch and Núria Sebastián-Gallés

Departament de Psicologia Bàsica, Universitat de
Barcelona, Barcelona, Spain

— Authors' copy —

This paper was published in *Cognition*, 64, (1997) B9-B17

It is well attested that we perceive speech through the filter of our native language: a classic example is that of Japanese listeners who cannot discriminate between the American /l/ and /r/ and identify both as their own /r/ phoneme (Goto, H., 1971. *Neuropsychologia* 9, 317-323.). Studies in the laboratory have shown, however, that perception of non-native speech sounds can be learned through training (Lively, S.E., Pisoni, D.B., Yamada, R.A., Tohkura, Y.I., Yamada, T., 1994. *Journal of the Acoustical Society of America* 96 (4), 2076-2087). This is consistent with neurophysiological evidence showing considerable experience-dependent plasticity in the brain at the first levels of sensory processing (Edeline, J.-M., Weinberger, N.M., 1993. *Behavioral Neuroscience* 107, 82103; Merzenich, M.M., Sameshima, K., 1993. *Current Opinion in Neurobiology* 3, 187-196; Weinberger, N.M., 1993. *Current Opinion in Neurobiology* 3, 577-579; Kraus, N., McGee, T., Carrel, T.D., King, C., Tremblay, K., Nicol, T., 1995. *Journal of Cognitive Neuroscience* 7 (1), 25-32). Outside of the laboratory, however, the situation seems to differ: we here report a study involving Spanish-Catalan bilingual subjects who have had the best opportunities to learn a new contrast but did not do it. Our study demonstrates a striking lack of behavioral plasticity: early and extensive exposure to a second language is not sufficient to attain the ultimate phonological competence of native speakers.

Introduction

Languages differ in the sounds they use to make up words and there is ample evidence that these differences have consequences for perceptual processing (e.g., Goto, 1971; Werker and Tees, 1984). More generally, research has shown that listeners are attuned to the properties of the speech signal which are most relevant for their language (e.g., Otake et al., 1993, 1996; Dupoux et al., 1997). Data exist that show that this tuning actually takes place very early, in the first year of life (see Kuhl, 1994; Werker, 1994 for reviews). An important issue arises for listeners who grew up in surroundings where two (or more) languages are spoken: can they perfectly master the two sound systems, and switch from one to

the other? Or do they use only one speech 'decoder' to process utterances from both languages? (Cutler et al., 1989, 1992; Mack, 1989).

When bilinguals' performance is assessed in production, it is clear that the later the start of exposure to the second language, the more noticeable is the 'foreign' accent (Oyama, 1976). For example, a 40-year-old immigrant who has spent the last 20 years in the US is more likely to speak with a strong accent than is an 8 year-old child who has lived in the US for 2 years. More precisely, most studies have shown that a non-native accent can be detected when the age of acquisition was above 6 years (for a comprehensive review, see Long, 1990). In a recent study, the foreign accent of 240 Italians immigrants was rated by native English listeners: the mean ratings decreased gradually with the age of arrival, and the most discerning listener could already detect an accent in the subgroup of Italians who had arrived before 4 years of age (Flege et al., 1995).

One possible locus for a non-native accent in pronunciation may be a loss of flexibility in learning new motor programs. If so, one may potentially expect more flexibility in perception. However, an effect of age of acquisition is also evident in perception: late learners are typically less sensitive to language-specific phonetic properties than early learners although less is known about the time schedule of the loss of plasticity in perception than in production. In

Address for correspondence: Dr. C. Pallier, Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS-EHESS, 54 bd Raspail, 75006 Paris, France. Tel.: +33 1 49542276; fax: +33 1 45449835; e-mail: pallier@lscp.ehess.fr This research was supported by grants from the Fyssen foundation, the French Ministère des Affaires Étrangères (programme Lavoisier), the Human Capital Mobility Program of the UE, and the Spanish Ministerio de Educación y Ciencia contract PB94926. We thank Anne Christophe, Anne Cutler, Emmanuel Dupoux and Jacques Mehler for useful comments on a previous version of this manuscript. This copy was typeset by Christophe Pallier using L^AT_EX and the apa package written by A. Protopapas

particular, it remains debated whether production loses flexibility before perception (Goto, 1971; Yeni-Komshian et al., 1973; Sheldon and Strange, 1982). The most recent data are consistent with the hypothesis that non-natives resemble native speakers more closely in perceiving than in producing some sounds (Flege, 1993). Moreover, to our knowledge, no previous published study has assessed the perceptual abilities of subjects who acquired their second language before the age of 6 years.

In the present study we have examined bilinguals who have been raised in a highly bilingual society and who have started learning their second language before 6 years of age. This has permitted us to avoid a difficulty inherent with studies on second language acquisition that have used immigrant subjects and looked for a correlation between performance and the age of arrival in the new country: comparing immigrant and native speakers creates notorious problems if one tries to equate the two groups of subjects from such populations. Late learners, for example, may simply have had less time or motivation to study the new language; in particular, they certainly have spent less time in school. In contrast, the native and non-native speakers that we have compared in the present study were all educated through the same schooling system.

Our subjects were from Catalonia, a region in the northern part of Spain. All were highly proficient speakers of Spanish and Catalan; indeed, in Barcelona, both languages are widely used in everyday life. Until the mid-seventies Spanish was the only official language and Catalan was mostly spoken in private circles, but since then Catalan has been fostered by a strong linguistic policy: for example, two of the four public TV channels broadcast exclusively in Catalan; in kindergarten, the use of Catalan dominates and in public primary schools children first learn to write Catalan before learning Spanish. At the university, most lectures are given in Catalan. As a result most young people are very good bilinguals and, in most places, it is very common to hear people switch from one language to the other in the middle of conversation.

Catalan and Spanish are both Romance languages and have many similarities, but their pronunciations differ significantly. In particular, Catalan has two mid vowels with different heights, one high [e] and one low [ɛ], while Spanish has only one [e] phoneme (which is more open than the Catalan [e]). The [e]-[ɛ] contrast is used to distinguish between common words in Catalan, e.g., [te] (take) and [tɛ] (tea), [pera] (Peter) and [pɛera] (pear). Our study focused on this [e]-[ɛ] contrast and assessed its perception by forty bilinguals with different backgrounds: half of our subjects had Spanish-speaking parents, and the other half Catalan-speaking parents. Thus, the latter were exposed to Catalan since birth, whereas the former were exposed to Spanish first, with exposure to Catalan starting in kindergarten or in primary school (at the latest at 6 years of age). To take part in the experiment, the subjects, who were all students at

the university of Barcelona, had to be born and have spent all their lives in Barcelona or its immediate vicinity (to avoid dialectal differences); moreover, only students who had been exposed to Catalan intensively since age 6 or earlier (range = 16; mean age = 4.35) were considered for inclusion in the experiment.

Stimulus preparation

We synthesized a series of seven vowel stimuli (s_1 - s_7) along a continuum from [e] to [ɛ] using Klatt's synthesizer (Klatt, 1980); they differed only in the frequency of the first formant that increased from 404 Hz (s_1) to 584 Hz (s_7) in steps of 30 Hz. For all stimuli, F2 was set to 2084 Hz, the duration to 200 ms, and a slowly declining pitch was imposed. These stimuli were employed in three tests administered to each subject: a classification task, an AX discrimination task and a typicality judgment task.

Experiment 1: classification along the /e-/ɛ/ continuum

In the classification task, the forty subjects were presented with a list in which each stimulus (s_1 - s_7) appeared nine times, in a randomized order; the task was to decide, after each stimulus, whether it sounded more like the first vowel in the Catalan word 'Pere' (Peter, pronounced [pera]), or like the first vowel in the Catalan word 'pera' (pear, pronounced [pɛra]); the two words, which are common in Catalan, were presented in written format in the instructions. Stimulus presentation and response recording was under the control of a computer. Fig. 1 displays the percentage of 'pera' ([pɛra]) responses for Spanish-born and Catalan-born subjects. In a 2×7 ANOVA performed on these scores, the factors 'Language' and 'Stimulus' interacted significantly ($F(6, 228) = 9.3; p < 0.001$). T-tests comparing the two 'Language' groups for each stimulus revealed significant differences on all but one (s_4).

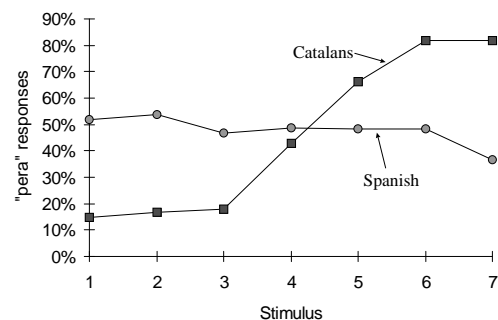


Figure 1. Categorization of the /e//ɛ/ continuum

Catalan-born subjects were readily able to map the stimuli onto the two words ‘Pere’ and ‘pera’, while the group of Spanish-born subjects produced an essentially flat response curve. Inspection of individual data revealed that there was great variability in the Spanish group: some subjects performed in a random fashion, others performed like most of the Catalan subjects and others systematically inverted the labels¹. One may interpret the data by suggesting that Spanish-born subjects simply are uncertain about the first vowels in the words ‘Pere’ and ‘pera’ (though these are quite common words). Their perception of vowel sounds would then not necessarily differ from the native Catalans. If such is the case, we would expect the two groups to behave similarly in the discrimination task to be used in the next experiment.

Experiment 2: AX discrimination experiment

In the discrimination task, subjects were presented with pairs of stimuli drawn from the s_1 - s_7 continuum. The randomized list contained seven ‘same’ pairs, repeated four times each, and five ‘different’ pairs, repeated eight times each. After each pair, subjects had to indicate by pressing a button whether the stimuli sounded physically identical or different. Fig.2 displays the rates of ‘different’ responses for the pairs that were actually different. A 2×5 ANOVA revealed a significant interaction between Language and Pair ($F(4, 152) = 4.9; p < 0.001$), and planned comparisons of the two groups restricted to each pair revealed that Catalan and Spanish groups differed significantly on the ‘ s_3 - s_5 ’ pair ($F(1, 38) = 5; p < 0.05$) and on the ‘ s_5 - s_7 ’ pair ($F(1, 38) = 7.7; p < 0.01$). Only the Catalan-born subjects show a peak in their discrimination performance for the s_3 - s_5 pair; this effect suggests that they partition the continuum into two phonemic categories separated by a boundary falling between s_3 and s_5 . Indeed, heightened sensitivity at a phoneme boundary is the hallmark of phonemic categorization (Repp, 1984)². Moreover, the Catalan group’s performance drops on the last pair. One potential explanation is that these stimuli are in the neighborhood of the Catalan /e/ phoneme prototype, as it is known that vowel discrimination

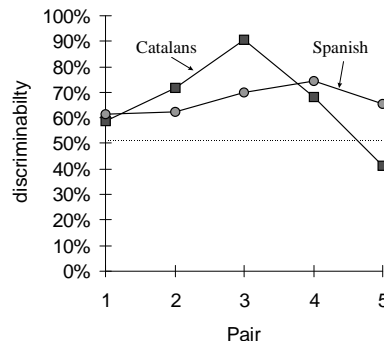


Figure 2. Discrimination of vowel pairs along the /e/-/ɛ/ continuum

is typically more difficult in the vicinity of a prototype (Kuhl, 1994). This remains a tentative explanation because one may wonder why the same effect is not observed at the other end of the continuum, around the /e/ prototype. This experiment reveals that the perception of the vowel continuum differs between the Catalan and the Spanish groups: it suggests that only those speakers who have been exposed to Catalan since birth possess two phonemic categories along the continuum. In the last experiment, we tried to gather more information about the internal representation of the vowel categories for the two populations.

Experiment 3: typicality judgments

Not all tokens of a phoneme sound like equally good exemplars to native speakers. It is possible to obtain more information about subjects’ representation of a phoneme category just by asking them to rate the typicality of each token (Miller, 1994). In this task, subjects were asked to rate the ‘category-goodness’ of each of the seven stimuli, relative to a Spanish [e], a Catalan [e] or a Catalan [ɛ]. For each trial, one of three words exemplifying a Spanish /e/ (mesa, ‘table’ in Spanish), a Catalan /e/ (Pere, ‘Peter’ in Catalan) or a Catalan /ɛ/ (pera, ‘pear’ in Catalan) was presented on the computer’s screen for 1 s; subjects were asked to form a ‘mental image’ of the first vowel in the word. Then, after the written word had been cleared, one of the seven vowel stimuli was played; subjects then had to give a score assessing the fit of the stimulus with the ‘imagined’ vowel (from 0 = ‘very bad’ to 5 = ‘very good’). The results are displayed on Fig.3. Data were analyzed, separately for each word, in 2×7 ANOVAs with the factors Language and Stimulus. The interaction between Language and Stimulus was significant for the Spanish [e] (mesa, $F(6, 228) = 4.9; p < 0.001$), and the Catalan [e] (Pere, $F(6, 628) = 4.9; p < 0.001$) but not for the Catalan [ɛ] (pera, $F(6, 228) = 1.2; p = 0.3$). Catalan-born subjects produced the ‘expected’ responses for the two

¹Individual data are available on the Internet from the homepage of the first author, Christophe Pallier (currently at <http://www.ehess.fr/centres/lscp/persons/pallier/articles.htm>).

²It is commonly held that consonant continua but not vowel continua are perceived categorically. Yet, according to (Repp, 1984p. 260): “In the earliest discrimination study, (Fry et al., 1962) found no discrimination peak at category boundary, but this was probably due to a ceiling effect, coupled with the use of imperfectly controlled stimuli. Most later studies (e.g., Fujisaki and Kawashima, 1969; Fujisaki and Kawashima, 1970; Pisoni, 1971; Stevens et al., 1969) have found clear peaks on vowel continua, so there is good reason to believe that there is a phonetic component in vowel discrimination”.

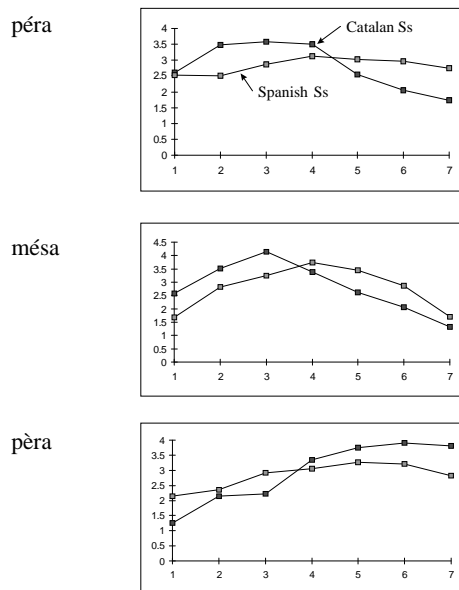


Figure 3. Typicality judgments

Catalan vowels: they preferred the stimuli in lowest F1 range for [péra] and the larger F1 range for [pèra]. Their behavior on [mésa] is interesting: their response curve is shifted towards the low F1 relative to the one of the Spanish group, and overlaps very much the [péra] curve. This suggests that native Catalan speakers conflate the Spanish and the Catalan [e] into one category. The behavior of the Spanish-born subjects is also interesting: they do not exhibit the same behavior on the three vowels. Their responses are flat for [péra], and show a trend toward the large F1 for [pèra]. These data show that Spanish-born subjects are not totally 'unaware' of the existence of different vowels in Catalan. However, they are a long way from performing like the group with Catalan as a first language.

Conclusion

Our results clearly document that Spanish-Catalan bilinguals who have learned Catalan from a very early age do not master the contrast between the Catalan [e] and [ɛ] phonemes. This extends previous studies of second language acquisition which have shown a loss of plasticity with age: the later a language is learnt, the less proficient people are (Oyama, 1976; Flege et al., 1995): it now seems that even early exposure (during the most sensitive period) is not sufficient. After exposure to Spanish leads to the formation of one [e] category, it appears to be difficult for the brain to learn two new phonetic categories which overlap with this one (see also Best et al., 1988). Similarly, Catalan-born sub-

jects who have first learned the two categories of their language apparently did not learn a third category when exposed to Spanish.

A caveat may be in order here: our results concern groups of subjects. It is not the case that two categories of subjects could be neatly distinguished a posteriori in the data. In other words, there was an overlap between the performances of the Spanish-born and the Catalan-born subjects. In part this is inevitable because of the noise in the responses we obtained, but this may also reflect a Catalan-like performance of some of the subjects in the Spanish group. Our data does not license the conclusion that none of the Spanish subjects learned the Catalan vowels. Spanish people living in Barcelona display a continuum of performance in Catalan and the best of them are not distinguished from natives with our methods. Nevertheless, it remains remarkable that many people exposed very early to a second language do not perform like native speakers.

It is clearly not the case that the bilinguals we tested possessed two independent phonetic spaces between which they might have easily switched. Moreover, our data provide strong evidence that the sound system they learned is very biased by the statistics of the first language they acquired. It would be interesting to investigate individuals who have been adopted very early (around 36 years of age) in a new country and who apparently have 'forgotten' their native language: has the second language acquired totally erased the traces of the first? This is a question for future research.

References

- Best, C.T., McRoberts, G.W., Sithole, N.M., 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.
- Cutler, A., Mehler, J., Norris, D., Segui, J., 1989. Limits on bilingualism. *Nature* 340, 229-230.
- Cutler, A., Mehler, J., Norris, D.G., Segui, J., 1992. The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24, 381-410.
- Dupoux, E., Pallier, C., Sebastian, N., Mehler, J., 1997. A distressing 'deafness' in French?. *Journal of Memory and Language* 36, 406-421.
- Edeline, J.-M., Weinberger, N.M., 1993. Receptive field plasticity in the auditory cortex during frequency discrimination training: selective retuning independent of task difficulty. *Behavioral Neuroscience* 107, 82-103.
- Flege, J.E., 1993. Production and perception of a novel second-language phonetic contrast. *Journal of the Acoustical Society of America* 93 (3), 1589-1608.

- Flege, J.E., Munro, M.J., MacKay, I.R.A., 1995. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16 (1), 1-26.
- Fry, D.B., Abramson, A.S., Eimas, P.D., Liberman, A.M., 1962. The identification and discrimination of synthetic vowels. *Language and Speech* 5, 171-189.
- Fujisaki, H., Kawashima, T., 1969. On the modes and mechanisms of speech perception (annual report of the Engineering Research Institute 28). Faculty of Engineering, University of Tokyo, Tokyo.
- Fujisaki, H., Kawashima, T., 1970. Some experiments on speech perception and a model for the perceptual mechanism (annual report of the Engineering Research Institute 29). Faculty of Engineering, University of Tokyo, Tokyo.
- Goto, H., 1971. Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia* 9, 317-323.
- Klatt, D.H., 1980. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America* 67 (3), 971-995.
- Kraus, N., McGee, T., Carrel, T.D., King, C., Tremblay, K., Nicol, T., 1995. Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience* 7 (1), 25-32.
- Kuhl, P.K., 1994. Learning and representation in speech and language. *Current Opinion in Neurobiology* 4, 812-822.
- Lively, S.E., Pisoni, D.B., Yamada, R.A., Tohkura, Y.I., Yamada, T., 1994. Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America* 96 (4), 2076-2087.
- Long, M., 1990. Maturation constraints on language development. *Studies in Second Language Acquisition* 12, 251-285.
- Mack, M., 1989. Consonant and vowel perception and production: early English-French bilinguals and English monolinguals. *Perception and Psychophysics* 46 (2), 187-200.
- Merzenich, M.M., Sameshima, K., 1993. Cortical plasticity and memory. *Current Opinion in Neurobiology* 3, 187-196.
- Miller, J.L., 1994. On the internal structure of phonetic categories: a progress report. *Cognition* 50, 271-285.
- Otake, T., Hatano, G., Cutler, A., Mehler, J., 1993. Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* 32, 258-278.
- Otake, T., Hatano, G., Yoneyama, K., 1996. Speech segmentation by Japanese listeners. In: Otake, T. and Cutler, A. (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies*. Mouton de Gruyter, Berlin, pp. 183-201.
- Oyama, S., 1976. A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5 (3), 261-283.
- Pisoni, D.B., 1971. *On the nature of categorical perception of speech sounds*. Unpublished Ph.D. thesis, University of Michigan.
- Repp, B.H., 1984. Categorical perception: issues, methods, findings. In: Lass, N.J. (Ed.), *Speech and Language: Advances in Basic Research and Practice*, Vol. 10. Academic Press, New York, pp. 243-335.
- Sheldon, A., Strange, W., 1982. The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3, 243-261.
- Stevens, K., Liberman, A., Studdert-Kennedy, M., O'hman, S., 1969. Cross language study of vowel perception. *Language and Speech* 12, 1-23.
- Weinberger, N.M., 1993. Learning-induced changes of auditory receptive fields. *Current Opinion in Neurobiology* 3, 577-579.
- Werker, J.F., 1994. Cross-language speech perception: developmental change does not involve loss. In Goodman, J.C., Nusbaum, H.C. (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. MIT Press, Cambridge, MA, pp. 93-120.
- Werker, J.F., Tees, R.C., 1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866-1878.
- Yeni-Komshian, G.H., Caramazza, A., Zurif, E.B., 1973. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54, 421-428.