

SPEECH PERCEPTION IN L2

Silvia Carmen Barreiro Bilbao. Universidad Nacional de Educación a Distancia (Madrid)

ABSTRACT

Theories of speech perception provide a theoretical framework within which to conduct experiments and test hypotheses. These theories should not only show the relation and degree of integration between different levels of information but should incorporate mechanisms to deal with developmental change or the effects of the linguistic environment on speech perception. The relevance of these factors is particularly important when concerning the perceptual difficulties encountered in the acquisition of second languages. The present paper is a review of the main current theories of L2 Phonetics which try to predict (and explain) the relative perceptual difficulty of non-native phonetic categories.

1. INTRODUCTION

The sounds of a language are organised into contrastive units (phonemes) which serve to distinguish words. But the level of speech organisation, that is, the way speech is systematically organised to convey meaning, differs from one language to another. As children we learn the relevant distinctions of our mother tongue (L1) and when we start to speak a second language (L2), we tend to produce sounds that we are used to making in our first language. This transfer from the L1 sound system to that new language gives 'foreign accent' to the pronunciation.

The process of L2 acquisition not only implies speech production but also speech perception, and it seems that there is a complex relationship between these two aspects (see J. Llisterri, 1995, for a review of the different issues concerning this relationship). The complexity is not only due to the theoretical issue of which goes first in the process of L2 acquisition (i.e., does production precede perception or vice-versa?), but also in terms of the many factors that should be considered when describing the relationship.

Leaving aside the theoretical discussions concerning the nature of the connection of these two aspects of L2 acquisition, in the present work we are going to focus on the perceptual side only (see Klatt, 1989, to get a general idea of how this complex mechanism is normally described theoretically). We believe that until we hear L2 sounds properly any attempt to reproduce them will be unsuccessful.

L2 speech perception will be discussed in terms of the problems encountered by non-native speakers and the variables that affect their perception in the first part of this paper. Current models of L2 Phonetics that aim to explain L2 perception process will be described

in the second part. From these models, hypotheses can be made about what should be easy or difficult to acquire for a L1/L2 pair, as we will explain.

2. L2 SPEECH PERCEPTION: PROBLEMS ENCOUNTERED AND VARIABLES

It is not easy to predict the perceptual difficulties that will be encountered by a L2 learner because of the number of problems and variables to take into account (Strange, 1995). A L2 learner will have perceptual difficulties when trying to differentiate speech contrasts that are not functionally distinctive in the first or mother tongue. As an example, Spanish listeners will have problems with the English contrast /s/ vs. /z/ as it is not phonemic in Spanish.

Also, they will find it difficult to differentiate phonetic categories that are phonemic in the native language, but differ in the phonetic realisation in L2. For instance, English listeners will have difficulty in differentiating the Spanish contrast /p, b/ due to differences in *VOT* and lack of aspiration in the Spanish voiceless sound.

With regard to the variables which interact to determine perceptual difficulty in the acquisition of non-native contrasts, Ocke-Schwen Bohn (1995) classified them into those which the researcher can *manipulate* (task variables), and those that the researchers *select* for analysis and study. The latter in turn can be divided into two groups: (i) Subject variables “which define (...) what the learner brings to the task of perceptually organising non-native contrasts” (84). (ii) “Contrast variables “which define (...) what the learner is trying to organise perceptually” (84).

The analysis of the *subject* variables can be summarised as follows:

a) L1 background: The difficulty in the perception of non-native contrasts is intimately related to the perceptual differentiation of phonetic contrasts in L1 and the relationship between L1 and L2 (Lisker & Abramson, 1970). Therefore, the influence of L1 will be stronger in some contrasts than others (Polka, 1991). Nevertheless, it is possible to minimise its effects, if there are variables, such as ‘L2 experience’, interacting with L1 (MacKain et al., 1981).

b) L2 experience: The immersion in a second language will alter the mechanisms underlying speech perception, and may even induce L2 learners to reorganise their linguistic world (Bohn and Fledge, 1990), “although several years of experience do not guarantee that L2 learners’ perception will become completely native-like” (Best & Strange, 1992: 327).

c) Age of the learner: This is a ‘key’ variable especially for those theories supporting the “critical period” of language acquisition, starting with Lenneberg in 1967. Since then, there has been a lot of work done on this issue. We can mention the most recent ones: (i) Mayo et al. (1997), who studied the effect of age of L2 acquisition and perception of speech in quiet and noise. They worked with English monolinguals, and early and late bilinguals (Mexican Spanish and English); and (ii) Fledge (1998), who analysed other variables in conjunction with age (such as how much the native language is used, and how long L2 has been spoken).

Other factors such as “attitudes towards L2”, “culture associated with L2”, and “motivation to learn L2” do not seem to contribute to the ability to differentiate non-native contrasts.

Contrast variables can be classified into two groups:

a) Type of contrast: some non-native contrasts are more difficult to differentiate

perceptually than others. For example, a total lack of experience in L1 with regard to ‘voicing’ contrasts would imply a great perceptual difficulty in differentiating L2 phonemic categories based on ‘voicing’ differences. On the other hand, if the L1 language uses ‘voicing’ to differentiate contrastive units, L2 learners would find it easier to acquire new non-native voicing contrasts. From the extensive work which has been done on this subject it is worth mentioning Lasky et al. (1975), for voicing contrasts, Werker et al. (1985), for consonantal place of articulation, or Polka (1995) and Fledge (1998), for vowel perception.

b) Type of cue: L2 learners seem to weight differently the information available in the signal from the way L1 learners do in making a linguistic decision. There has been a lot of work done in this area: Strange and Jenkins (1978), Yamada and Tohkura (1992), Flege (1984), Bohn (1995), Bohn & Flege (1990), and Hazan & Boulakia (1993). Studies with Spanish would include Flege and Eefting (1987), and L. Williams (1977), for Spanish-English bilinguals.

The issue of ‘cue multiplicity’ is extremely important. The fact that phonemic contrasts are signalled by several acoustic cues implies that there is redundant information in the signal. Therefore, even in those cases when the signal is degraded (speech through telephone, noisy environment, etc.), we may still be able to recognise the sounds based on the presence of cue only. Furthermore, it raises the issue of ‘perceptual weighting’: listeners may make use of different acoustic cues present in the signal when trying to differentiate phonetic contrasts (Hazan & Rosen, 1991; Hazan & Shi, 1992; Barreiro, 1998). This different perceptual behaviour can also be applied to L2 learners, as mentioned above.

Despite the amount of work on L2 perception, further work needs to be done (Bohn, 1995) to find out what makes non-native listeners use cues that are not used by native listeners, and what variables interact with them.

3. MODELS OF SECOND LANGUAGE PHONETICS

The fact that languages have different phonological systems and make use of different phonemic categories has been a matter of concern in theories of L2 Phonetics. A L2 learner has to learn contrasts that are different from those learnt in L1. Furthermore, not all non-native phonetic contrasts are equally difficult for L2 learners, as we have stated above.

All the current theories in L2 speech perception have tried to predict (and explain) the relative perceptual difficulty of those non-native phonetic categories. The four models that will be explained in this section can be considered the main current theories of L2 Phonetics, according to Strange.

3.1. Perceptual Assimilation Model (PAM)

According to this model, we can predict the perceptual difficulties in differentiating a non-native phonetic contrast if we take into account the degree of its articulatory-phonetic (gestural) similarity to native *gestural constellations* (Best & Strange, 1992). Therefore, it seems that listeners do not hear acoustic speech signals, but rather the phonetic gestures that produced that signal.

If the contrasting sounds are both assimilated as good exemplars of a single L1 category, it will be very difficult for the listener to perceive a difference. Consequently, there will be no discrimination. However, if they vary in goodness of fit to a single L1 category,

then the perceptual discrimination will be much easier, as they will be associated with two different native categories. Finally, if the non-native phones are very dissimilar to native phonetic gestures, they may be assimilated as ‘uncategorizable’ or even as ‘non-speech’ sounds.

This model assumes the concept of ‘critical period’, supported by the *Maturational State Theory*: A second language is spoken without a ‘foreign accent’ if it is learnt before the end of the critical period for language acquisition.

3.2. Speech Learning Model (SLM)

Flege’s model (1995) supports the idea that there is a clear interaction between L1 and L2. L2 sounds or segments which are similar (in acoustic-phonetic terms) to L1 sounds will be assimilated to native categories. On the contrary, if there is a great difference perceptually, new L2 categories will be established over time. In those cases in which the L1 and L2 sounds are not extremely different, a single category subsuming both sounds will be created, and there will be a clear accented production of the L2 sounds. Furthermore, the production of those L1 sounds can also be affected.

This model assumes the existence of a ‘sensitivity period’ after which it is more difficult to acquire non-native phonetic contrasts. They believe that there is no evidence that the ability to acquire a new phonetic contrast diminishes after the critical period of maturation.

3.3. Perceptual Learning Model (PLM)

Pisoni and his colleagues (1994) proposed a theory that states that infants are not born with a clear ability to discriminate speech sounds, although they have good auditory sensitivity. It is only by early experience that they will develop the perceptual ability to discriminate phonetic contrasts. As Pisoni et al. said: “its development may depend on a process of *induction* based on specific early experiences the organism has in the environment” (124). Therefore, their discrimination will be highly dependent on the environmental language and will be initially limited to those phonetic contrasts that are present in the native language.

They carried out some experiments with adults that showed the following: (1) the apparent developmental loss brought by acquiring a language is a change in selective attention rather than a true sensory-based loss. (2) Our ability to discriminate non-native phonemic differences are not permanently modified or lost in adults. (3) Perception is affected by linguistic experience.

Their exemplar-based model proposes that listeners store detailed information about individual phonetic segments, which includes information about the speaker and the context. Native phonetic categories are represented as clusters of exemplars that share certain critical (acoustic) parameters, while varying on other, non-critical characteristics.

The success of L2 perceptual learning will depend on the degree of reorganising those phonetic categories, with the formation of (new) equivalent clusters.

3.4. Native Language Magnet Model (NLMM)

In the early nineties Kuhl proposed a new model of speech perceptual development based on the results of a series of experiments carried out with children and adults (Kuhl,

1991b). Two fundamental concepts need to be explained, the term 'prototype' and the 'perceptual magnet effect'. A 'prototype' is the best, most typical representative of a category, which is stored in the long-term memory, and it is used as a standard of comparison. This prototype acts, according to Kuhl and her colleagues (Kuhl & Iverson, 1995), as an attractor on the surrounding sounds in the same speech category. It is difficult to differentiate the prototype from the sounds that are acoustically similar. Incoming speech-sounds are identified by reference to the category prototype acting as a perceptual magnet, because they are completely or relatively indistinguishable from it. Poorer examples of the speech-sound category ('non-prototypes') do not function as perceptual magnets.

Therefore, according to this model, Magnet Effects precede changes in infants' perception of foreign non-native contrasts (Kuhl, 1995). Furthermore, learning a mother language results in alterations of the underlying perceptual mechanisms that affect the processing of language from that time onwards.

Kuhl's theory tries to explain the state of speech perception at birth and the different changes suffered during the first year of life (1991a, 1992). Infants are born with natural auditory boundaries between sounds that exist in the languages of the world. Those boundaries are not determined by language experience but are due to auditory processing mechanisms. They can discriminate any contrast independently from their native language. With time, their perceptual boundaries change as a result of language exposure. Finally, by the end of the first year, prototypes act as perceptual magnets leading to the failure of discrimination of foreign contrasts.

Linguistic experience in different cultures results in magnet effects that differ in number and location for infants growing up listening to the different languages (based on Kuhl & Iverson, 1995). As adults, we may have perceptual problems in L2 units depending on its proximity to our native-language magnet: "The nearer it is to a magnet, the more it will be assimilated to the native-language category, making it indistinguishable from the native-language sound" (Kuhl & Iverson, 1995:143).

Therefore, acquiring a second language will cause the reorganisation of our phonetic perceptual space around newly established prototypes.

The ideas behind this model of speech perception are still evolving. In fact, Dr. Sarah Barrett (1997) has designed a new theory of Speech Perception, called the *Prototype Conversion Model (PCM)* which expands Kuhl's *NLM* proposing a new use of the idea of phonetic prototype to explain children's perceptual development of speech. This new model, such as the idea that perceptual magnets are context-independent and limited to speech (Barrett, 1999) is questioning some of Kuhl's claims.

4. CONCLUDING REMARKS

Theories of speech perception provide a theoretical framework within which to conduct experiments and test hypotheses. This theoretical framework can be applied in the field of speech technology or in speech and language therapy, for example. Furthermore, these theories should not only show the relation and degree of integration between different levels of information (acoustic, phonological, semantic, etc.) but should incorporate mechanisms to deal with developmental change or the effects of the linguistic environment on speech perception. The relevance of these factors is particularly important when concerning the perceptual difficulties encountered in the acquisition of a second language.

Those difficulties with certain speech contrasts and variables, such as L1 background, L2 experience, etc., were described in the first section, following Ocke-Schwen Bohn's ideas.

In the second section, we presented four models in an effort to predict the relative perceptual difficulty of non-native phonetic categories. These models make claims about the nature of the stored representations of L1 phonetic categories. Nevertheless, there were differences among them:

The first two models, *Perceptual Assimilation Model* and *Speech Learning Model*, differ in the emphasis placed on acoustic versus articulatory specification of phonetic similarity. However, both consider appropriate to analyse at the level of phonetic segments independently from the context rather than the more abstract phonemes or distinctive features of traditional analysis. On the other hand, the other two models, *Perceptual Learning Model* and *Native Language Magnet Model*, are more concerned with the nature of organisation of phonetic categories and how it affects the perception of native and non-native phonetic segments (Strange, 1995).

Besides, the implications of the different approaches varied greatly in relation to a deeper issue, that is, if the underlying neural mechanisms used in speech perception that become finely tuned to the distinctive contrasts used in the native language represent a reversible or non-reversible process.

If it is non-reversible, it will be impossible to learn new phonetic contrasts past a critical period (age). However, if the learning of phonetic contrasts of a language is the result of selective attention, through training (Jamieson, 1995), it might be possible to redirect attention towards new contrasts, (relevant for L2 acquisition, auditory training for deaf children with delayed speech development, etc.).

Nowadays, it seems that the second position is more broadly accepted: "perceptual difficulties are *not* due to a loss of sensory capabilities, but rather reflect *perceptual attunement* to phonetic information that is phonologically relevant in their native language" (Strange, 1995: 79). Support for this position comes from the results of the techniques applied for training in non-native speech contrasts (Jamieson, 1995; Barreiro, 1998). Our own pronunciation and perception are a good example of the success of those techniques.

REFERENCES

- Barreiro Bilbao, S.C. 1998. *Spanish Perception of English Fricative 'Voicing' and 'Place of Articulation' and the Effect of Cue-Enhancement on Their Perception*. Unpublished MSc dissertation. London: University College London.
- Barrett, S. 1997. *Prototypes in Speech Perception*. Unpublished doctoral dissertation. Cambridge: University of Cambridge.
- Barrett, S. 1999. "The perceptual magnet effect is not specific to speech prototypes: new evidence from music categories". *Speech, Hearing and Language: Work in Progress*. U.C.L 11. 1-16.
- Best, C. & W. Strange. 1992. "Effect of experience and phonetic factors on cross-language perception of approximants". *Journal of Phonetics* 20: 305-330.
- Bohn, O. 1995. "What determines the perceptual difficulty encountered in the acquisition of nonnative contrasts?". *ICPhS 95*, 4: 84-91.

- Bohn, O. & J. E. Flege. 1990. "Interlingual identification and the role of foreign language experience in L2 vowel perception". *Applied Psycholinguistics* 11: 303-328.
- Flege, J.E: 1984. "The effects of linguistic experience on Arabs' perception of the English /s/ vs. /z/ contrast". *Folia Linguist* 18: 117-138.
- Flege, J.E. 1995. "Second language learning: theories, findings and problems". *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Ed. W. Strange. Timonium MD: York Press. 215-24.
- Flege, J.E. 1998. "Second-language learning: the role of subject and phonetic variables". *ESCA-StiLL* 98: 1-8.
- Flege, J.E. & X. Eefting. 1987. "Production and perception of English stops by native Spanish speakers". *Journal of Phonetics* 15: 67-83.
- Hazan, V.L. & G. Boulakia. 1993. "Perception and production of a voicing contrast by French-English bilinguals". *Language and Speech* 36: 17-38.
- Hazan, V.L. & S. Rosen. 1991. "Individual variability in the perception of cues to place in initial stops". *Perception and Psychophysics* 49: 187-200.
- Hazan, V.L. & B. Shi. 1992. "Listener variability of synthetic and natural speech". *Speech, Hearing and Language: Work in Progress*. U.C.L. 6: 75-88.
- Jamieson, D.G. 1995. "Techniques for training different non-native speech contrasts". *ICPhS 95*, 4: 100-107.
- Klatt, D.H. 1989. "Review of selected models of speech perception". *Lexical Representation and Process*. Ed. Marslen-Wilson. MA: MIT Press. 145-56.
- Kuhl, P.K. 1995. "Mechanisms of developmental change in speech and language". *ICPhS 95*, 2: 132-139.
- Kuhl, P.K. 1991a. "Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech". *Plasticity of Development*. Eds. W.S. Brauth, W.S. Hall, & R.J. Dooling. MA: MIT Press. 73-106.
- Kuhl, P.K. 1991b. "Human adults and human infants show a 'Perceptual Magnet Effect' for the prototypes of speech categories, monkeys do not". *Perception & Psychophysics* 50: 93-107.
- Kuhl, P.K., K.A. Williams, F. Lacerda, K.N. Stevens & B. Lindblom. 1992. "Linguistic experience alters phonetic perception in infants by 6 months of age". *Science* 255: 606-608.
- Kuhl, P.K. 1993. "Early linguistic experience and phonetic perception: implications for theories of developmental speech perception". *Journal of Phonetics* 21: 125-139.
- Kuhl, P.K. & P. Iverson. 1995. "Linguistic experience and the 'Perceptual Magnet Effect'". *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Ed. W. Strange. Timonium MD: York Press. 121-153.
- Lasky, R.E., A. Syrdal-Lasky & R.E. Klein. 1975. "VOT discrimination by four to six and a half month old infants from Spanish environments". *Journal of Experimental Child Psychology* 20: 215-225.

- Lisker, L. & A.S. Abramson. 1970. "The voicing dimension: some experiments in comparative phonetics". *Proceedings 6th International Congress of Phonetic Sciences*: 563-567.
- Llisterri, J. 1995. "Relationships between speech production and speech perception in a second language". *ICPhS 95*, 4: 92-99.
- Mayo, L.H., M. Florentine & S. Buus. 1997. "Age of second-language acquisition and perception of speech in noise". *Journal of Speech, Language and Hearing* 40: 686-693.
- MacKain, K., C. Best & W. Strange. 1981. "Categorical perception of English /r/ and /l/ by Japanese bilinguals". *Applied Psycholinguistics* 2: 369-390.
- Pisoni, D.B., S.E. Lively & J.S. Logan. 1994. "Perceptual learning of nonnative speech contrasts: implications for theories of speech perception". *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Eds. J.C. Goodman & H.C. Nusbaum. Cambridge MA: MIT Press. Chapter 5.
- Polka, L. 1991. "Cross-language speech perception in adults: phonemic, phonetic, and acoustic contributions". *J.A.S.A.* 89: 2961-2977.
- Polka, L. 1995. "Linguistic influences in adult perception of non-native vowel contrasts". *J.A.S.A.* 97: 1286-1296.
- Strange, W. 1995. "Phonetics of second-language acquisition: past, present, future". *ICPhS 95*, 4: 76-83.
- Strange, W. & J.J. Jenkins. 1978. "The role of linguistics experience in the perception of speech". *Perception and Experience*. Eds. R.D. Walk & H.L. Pick. New York: Plenum Press. 125-169.
- Werker, J.F. & J.S. Logan. 1985. "Cross-language evidence for three factors in speech perception". *Perception and Psychophysics* 37: 35-44.
- Williams, L. 1977. "The perception of stop consonant voicing by Spanish-English bilinguals". *Perception and Psychophysics* 21 (4): 289-297.
- Yamada, R. & Y. Tohkura. 1992. "Perception and production of syllable-initial English /r/ and /l/ by native speakers of Japanese". *Proceedings of the International Conference on Spoken Language Processing*: 757-760.