RELATIONSHIPS BETWEEN SPEECH PRODUCTION AND SPEECH PERCEPTION IN A SECOND LANGUAGE

Joaquim Llisterri
Departament de Filologia Espanyola, Facultat de Filosofia I Lletres, Edifici B, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. Fax: +34.3.581.16.86; E-mail: joaquim.llisterri@cc.uab.es

ABSTRACT
This contribution presents a review of some of the findings reported in the literature concerning the relationship between production and perception of the sounds of a second language. Part of the experimental research and some applied work is summarized. It is concluded that a complex relationship exists between the production and perception of L2 sounds, and that many factors have to be taken into account in describing this interaction.

INTRODUCTION
Research concerned with the relationship between the production and perception of the sounds of a second language (L2) has addressed a problem that can be summarized as follows: Does production precede perception or, conversely, does perception precede production in the process of acquiring an L2? That is to say: can learners adequately pronounce sounds which are not well perceived, or is a good perception a prerequisite to accurate pronunciations? The answer to this question has not only theoretical implications regarding the process of L2 acquisition, but also practical consequences as far as the methodology used for teaching pronunciation is concerned.

Perception precedes production
As early as in 1931, Polivanov [1] claimed that the phonemic representations of a second language are perceived according to the system of the first language; although it is difficult to assess the validity of the data supplied by Polivanov (see Rochet [2] note 3), his remarks have been interpreted as supporting the hypothesis that difficulties in the production of the sounds of a second language arise from the influence of the L1 phonological structure on the perception of L2 sounds.

Production precedes perception
As Borrell [6] points out, it is a very common experience when learning an L2 that not all the sounds that are correctly perceived will be correctly produced. Similar observations have been made by Neufeld [7] and by Brière [8]. It seems then, that in certain cases, the production of L2 sounds might precede their perception.

In order to find evidence supporting one or both principles outlined below, we will review part of the literature concerned with the relationship between speech perception and speech production both in L2 and in bilinguals.

The review is divided in two parts: in the first one, the major findings of selected experimental studies concerning segmental elements will be summarized. The second part will concentrate on work carried out from a more pedagogical perspective, trying to establish the efficiency of pronunciation teaching methods that emphasize initial training either in production or in perception.

Studies addressing this topic in the first language will not be mentioned, but it is worth noting that "it is not clear how production and perception are linked in mature native speakers" [9] (p.51). The literature on L1 development will not be surveyed either, although undoubtedly it may shed some light into the issue.

Another field that should be taken into account but will not be discussed here is speech pathology. It should be reminded that the proponents of the verbo-tonal method have advocated that the situation of a deaf person, in the sense that lack of an adequate perception inhibits accurate productions. Thus, the notion of 'phonological deafness' has been applied to second language learners in this context [5].

EXPERIMENTAL EVIDENCE OF RELATIONSHIPS BETWEEN PRODUCTION AND PERCEPTION IN L2
The results that can be found in the literature will be presented separately for vowels and for consonants since, as will be shown later, it has been proposed that different classes of sounds may behave in a different way concerning the relationship between production and perception [9].

The production and perception of L2 vowels
The experimental studies concerning vowel production and perception have consistently shown a close link between those abilities in L2 learners.

The strength of this link has been emphasized, for example, by Barry [10]. After examining the correlation between pronunciation and perception of English vowels by German speakers, he concluded that "well-established perceptual categories are more likely to be accompanied by more acceptable production" (p.160), suggesting the possibility of using perceptual abilities to predict accuracy in production.

The perceptual bases of production errors have been also examined in a recent experimental study by Rochet [2]. Speakers of Canadian English and of Brazilian Portuguese were asked to imitate the French vowel /y/ and to label a continuum of synthetic high vowels as [i] or [u]. Production errors - systematic substitutions of French [y] by [u] in the case of English speakers, and by [i] in the case of Portuguese speakers - were correlated to labelling results in a perceptual test: as predicted by the hypothesis, vowels with a second formant in the range of French [y] were mostly identified as [u] by English speakers and as [i] by Portuguese speakers.

According to Rochet, these results support the notion that "accented pronunciations of L2 sounds by untrained speakers may be perceptually motivated"; they also seem to imply a good correlation between perception and production, the first taking precedence over the second.

However, other studies suggest that the perception-production link is of a more complex nature, and might be influenced by different factors. Elsendoorn [11] compared the results of an auditory task in which Dutch learners of English had to adjust the duration of vowels in English words with durational data obtained from words produced by the same group of subjects. The results showed an effect of the fortis or lenis character of the final obstruent in the perceptual task that was not present in the production data. Another interesting result of the experiment was that standard deviations of perceptual adjustments diminished "with growing familiarity with and knowledge of the English language" (p.131). This introduces a new dimension that increases the complexity of the production-perception link, namely the knowledge of L2.

Bohn & Flege [9] directly approached the influence of this factor by examining the production and perception of the English vowels /e/ and /æ/ in two groups of German learners of English: an experienced group - consisting of
individuals who had lived in the United States for at least five years - and an inexperienced one - formed by learners with a mean length of stay in the US of six months.

The results of the experiment revealed a clear difference between the two groups of speakers. The inexperienced German learners did not produce a contrast between the two vowels under consideration, but they succeeded in differentiating the vowels in a labelling task; however, they relied primarily in duration, a cue that does not strongly influence the labelling of the same continuum by native English speakers. Experienced German speakers of English did produce the contrast between the two vowels and were able to achieve better labelling results than the non-experienced ones, but they did not use durational and spectral cues in the same way that native English speakers did in the same labelling task. Bohn & Flege [9] also reported that while spectral differences in production were related to a dominance of spectral cues in perception, strong reliance on durational cues in perception implied small durational differences in production.

These results point out that more elements are to be taken into account when attempting to explain the relationship between production and perception in vowels: on the one hand, experience with the language seems to percep-tually vowels that they were not able to contrast in production [9], and the fact that both the Portuguese and the English speakers studied by Rochet [2] correctly imitated the French vowel [y] in roughly 50% of their French productions seems to imply that production abilities can not be completely inferred from perceptual ones and vice-versa. In fact, Bohn & Flege [9] found that experience might improve production, since "continued L2 contact enables L2 speakers to produce a new vowel contrast like native speakers of the L2"; but at the same time, they noted that "perception abilities for a new vowel contrast may lag behind even after several years of L2 experience" (p.52).

Moreover, even if discrimination was correct, Bohn & Flege convincingly argued that the strategies used by L2 learners (experienced or not) might not be the same as the ones used by native speakers. (See also Bohn, this volume, for an explanation of the results).

In summary, as far as vowels are concerned, there is a complex link between production and perception in L2 sounds. Although it seems that perception in general might precede production, direct inferences about pronunciation accuracy can not probably be made from perceptual abilities in a straightforward manner. Factors such as contextual dependency, nature of the acoustic cues involved in phonetic contrasts and the learner's familiarity with the L2 must as well be taken into account.

The production and perception of L2 consonants

In order to exemplify the relationship between production and perception of consonants in a second language, we will summarize some findings concerning two widely studied classes: liquids and stops.

The studies concerning the production and perception of liquids have concentrated on the distinction between English /t/ and /d/ by speakers of languages such as Japanese and Korean, which do not contrast these segments in their phonological system. An experiment was carried out by Sheldon & Strange [12] with Japanese speakers of English living in the United States. It was shown that the production of the English contrast between /t/ and /d/ was more accurate than the perception of natural utterances; the materials of the perceptual test comprised minimal pairs with /t/ and /d/ taken both from native productions and also from the subjects' own productions of the pairs.

The same findings were reported by Goto [13] in a previous experiment carried out in Japan; as concluded by Sheldon & Strange "at least for the contrast studied here, perceptual mastery of a foreign contrast does not necessarily precede adult learners' ability to produce acceptable tokens of the contrasting phonemes, and may, in fact, lag behind production mastery" (p.245).

Flage [5] discusses some of the facts that may have influenced the results reported by Sheldon & Strange: explicit articulatory training undergone by some of the speakers (see also the results of Catford & Pisoni's [14] experiment reported below), characteristics of the corpus - i.e., words in citation form - and the monitored nature of the situation. Sheldon & Strange themselves noted that the group of speakers had very specific characteristics, in the sense that they were at an advanced state of L2 acquisition.

Bohn, Gerber & Milsark [15] examined the relationship between perception and production of English /n/ and /l/ in Korean learners of English in an experiment that comprised production, identification, discrimination and a self-perception test. Synthesized stimuli were used for the identification and discrimination test. One of the main results obtained was that self-perception develops earlier and may be a prerequisite for accurate production.

The authors note that "the ability to make phonemic perceptual judgments in an /l/ - /n/ continuum that are similar to those of English speakers also seems to improve before production" (p. 516).

This results for Korean learners seem to be the most in partial disagreement with those reported by Sheldon & Strange [12] for Japanese speakers. However, the results obtained by Borden et al. have been reanalyzed by Sheldon [16]. The statistical treatment that Sheldon applied to the data allows for an interpretation of the results that is more coherent with the findings of Sheldon & Strange [12], since the idea that self-perception precedes production is not confirmed.

One of the important conclusions of Sheldon's reanalysis was that the relationship between production and perception depended on the amount of time spent in the USA by the Korean learners, so that "as the learner's time in the US increases the probability of occurrence of perception exceeding production decreases" (p.111). This would be in agreement with the fact that the speakers studied by Sheldon & Strange [12] were advanced learners.

The model proposed by Sheldon is not only corroborated by her interpretation of the Borden et al. data, but also by the fact that, as she hypothesizes, a functional perceptual level for an L2 might be enough for communication purposes, while heavily accented productions are socially less accepted, with the consequence that L2 speakers would have more pressure to improve production than perception. The same conclusions from an experiment concerned with /t/ and /d/ were reported by Sheldon [16]. The statistical treatment that Sheldon applied to the data allows for an interpretation of the results that is more coherent with the findings of Sheldon & Strange [12], since the idea that self-perception precedes production is not confirmed.

We may now examine some of the results reported for the production and perception of stop consonants in L2. Flage & Eefting [17] studied the production and perception of the /n/ - /l/ contrast by Dutch learners of English with different degrees of familiarity with the language. The results showed that
Dutch speakers were able to produce a substantial VOT difference between Dutch and English, indicating a good discrimination of the two languages. However, in perception tests, even the most proficient Dutch speakers showed only a small shift in the location of phoneme boundaries when identifying stops in a /d/ - /t/ synthetic continuum, which they were induced to perceive as Dutch or English by modifying the language settings of the percept. This seems to suggest that the distinction between the two languages in perception was not as clear as in production. The discrepancies between the cross-language shift in production and perception lead the authors to conclude that there is a disparity between production and perception.

The same disparity can be observed in a study that directly addressed the influence on production of adequate perceptual representations of the sounds of L2. Flege [18] examined vowel duration as a cue to voicing in English words ending with /t/ or /d/ produced and perceived by Chinese speakers of English; subjects were classified according to age of L2 acquisition and experience with the language.

The existence of a link between production and perception was revealed by the correlations between differences in perceived vowel duration and degree of foreign accent judged by native speakers of English; correlations between voicing effects in production and differences in vowel duration in perception were found too. Nevertheless, it appeared that "non-natives will resemble native speakers more closely in perceiving than in producing vowel duration differences" [18] (p.1605), a result that would be in agreement with the 'perception before production' hypothesis, specially in the case of experienced learners.

However, it can be concluded from the study of individual differences that, in adult learners, production is not limited or inhibited by the perceptual representations of the L2 sounds; the explanation proposed by Flege to account for small differences in vowel duration found in some of the subjects together with large effects in perception are based on problems with the timing of articulatory gestures and on a reduced sensitivity to vowel duration differences due to self-hearing (since a correlation was found between unsuccessful imitations and strong foreign accent).

The conclusions of the Borden et al. study [15] about the importance of self-perception in the development of production accuracy and the possibility of developing a near-native perceptual ability before reaching the same level in production seem to agree with some of Flege's findings.

Finally, it is hypothesized by Flege, that the presence of more experienced learners in the group of subjects would have lent support to the precedence of perception over production.

Another study on the perception and production of Italian stop consonants by Austrian German learners carried out by Grasberger [18] also supported the hypothesis that well-established perceptual categories do help accurate productions. He also suggested, as Barry [10] did for vowels, that perceptual tests might be a good tool to indicate production difficulties.

It can be seen again that, as far as consonants are concerned, it is not easy to establish a direct correlation between production and perception in an L2 although obviously some links do exist.

The nature of this link does not seem clear if we try to integrate the experimental results. On the one hand, Sheldon & Strange [12] showed that production can precede perception in advanced learners; the Dutch speakers of English examined by Flege & Eefting [17] presented a better differentiation of the two languages in production than in perception; also, some of the learners examined by Flege [18] showed larger production effects than perception effects. Sheldon's [16] hypothesis about the greater social pressure to improve production than perception could be a plausible explanation for these facts.

On the other hand, the findings of Borden et al. [15] seem to point out in the direction of a precedence of perception over production and the group results obtained by Flege [18] also seem to suggest a more native-like behaviour in perception than in production.

Finally, it should be noted that, as in the case of vowels, the age of L2 acquisition, the degree of exposure to the language, and the experience with L2 seem to be factors that affect the general correlation between production and perception.

Detection of foreign accent

Another useful contribution to the debate on the precedence of perception over production can be found in studies concerning foreign accent detection. Early work by Flege [20] will be summarized as an example.

Comparison of the performance of a group of Taiwanese subjects who had lived in the US for one year and another group who had lived there for five years showed that, in a foreign accent detection task, the experienced group was able to distinguish non-native from native speakers of English better than the non-native group. However, both groups were rated as having equally strong foreign accent by native English judges.

According to Flege, this shows a dichotomy between speech production and speech perception, and it can be convincingly argued that, for the subjects of the experiment, their ability to detect non-authentic productions was greater than their ability in production.

Flege, then, proposed that perception is more subject to improvement with time than production is. However, the conclusions of Bohn & Flege's study of vowels [9] suggesting that "perception abilities for a new vowel contrast may lag behind even after several years of L2 experience" (p.52) do not seem to agree with the earlier findings; Sheldon's explanation [16] based on the social need to improve production more than perception seems to be more coherent with the 'production precedes perception' hypothesis.

Production and perception of L2 sounds in bilingual speakers

To conclude this part of the review, some experimental studies concerning the relationship between production and perception in bilingual speakers will be presented.

Caramazza et al. [21] compared the voiced/unvoiced contrast in stops produced and perceived by Canadian French-English bilinguals with the results obtained from monolingual speakers. They found that "the bilingual subjects produced voicing distinctions which were clearly different for the two languages; and this disparity stands in marked contrast with the similarity of their perceptual functions in the two language 'modes'" (p.425); their conclusion was that bilingual speakers can better adapt their production than their perception in their non-dominant language. (See Hazan & Boulakia [22] for the effect of language dominance on bilingual's performance).

Similarly, in a study of the production and perception of English /t/ - /d/ and /l/ contrasts in early French-English bilinguals with English dominance, Mack [23] found evidence that "bilingual production can be more accurate than perception" (p.197).

The findings of both studies seem to point out towards a better differentiation between the two languages in production than in perception. The explanation suggested by Mack is similar to the one proposed by Sheldon [16] to account for the same trend in L2 speakers: the social consequences of non-native production are more important than those of non-native perception and, therefore, accurate productions are found whereas perception can be different from monolinguals, whenever comprehension is achieved.

PRODUCTION, PERCEPTION AND PRONUNCIATION TEACHING

Having reviewed some of the evidence found in the experimental literature, it might be useful to consider studies that have approached our topic from a perspective more oriented towards the teaching of pronunciation.

The techniques for training non-native contrasts will not be presented here (see Jamieson, this volume); instead, we will concentrate on work discussing the effects of training based on production vs. training based on perceptual strategies.

In a classical experiment, Catford & Pisoni [14] showed a superiority in production and perception of a set of so-called 'exotic sounds' in subjects who received an articulatory training compared to subjects that were trained...
with auditory techniques based on perceptual discrimination.

This led the authors to conclude that "what is effective in the teaching of sound production is the systematic development by small steps from known articulatory postures and movements to new and unknown ones" (p.477), implying, at the same time, that good production abilities may contribute to a better discrimination of L2 sounds.

The same conclusion seems to be supported by Weiss [24], reporting a previous experiment in which it was shown that training in pronunciation improved the discrimination abilities of a group of Chinese students of English. Another interesting result of this study was that a greater experience in a second language improves perception more than production (cf. [20] discussed earlier for a similar result).

It seems, then, that it can be provisionally concluded that training in production might help to improve production. However, according to the results provided by Rochet [2], the opposite also seems to hold true.

Rochet claimed that a significant improvement in the production of French voiceless stops by native speakers of Mandarin Chinese - from 30% to 19% of incorrect pronunciations - was found after perceptual training with synthetic Mandarin speech. Moreover, "the acquisition of L2, experience with the language, degree of exposure, and age of acquisition seem to play a major role in the interaction between production and perception in L2" (p.477). It does not seem possible to infer production abilities from perceptual ones and vice-versa.

Stage in the acquisition of L2, experience with the language, degree of exposure, and age of acquisition seem to play a major role in the interaction between production and perception in L2. The relation between production and perception might differ according to the class of sounds, to the acoustic and perceptual correlates of these classes and to contextual effects.

Social factors such as pressure to improve production may provide an explanation for cases in which production precedes perception.

Moreover, methodological problems have to be considered. First of all, Mack [23] has mentioned the inadequacy of comparing results from tests of speech production with results derived from speech perception tests, since there are important differences in the nature of the techniques used to assess these activities. Secondly, Bohn & Flege [9] point out those difficulties in defining the criteria used to evaluate accuracy in production and perception, which lead to difficulties in comparing different studies. Finally, it has to be reminded that most of the work carried out has concerned experimental studies in which highly controlled tasks and linguistic materials have been used.

It seems, in summary, that we have come a long way in the characterization of production and perception skills, but as a recent paper by Rochet [2] concludes, "the relationship between the perception of L2 speech sounds and their production by non-native speakers is still far from well understood".

ACKNOWLEDGMENT

Thanks are due to Ninon Font and Astrid Roig for their help in the survey on which this paper is based.

REFERENCES