The effects of two higher-level variables on the processing of an allophonic contrast in European Portuguese

JOÃO VELOSO
University of Porto (Portugal)

1. INTRODUCTION

1.1. Theoretical background

In this paper, our purpose is to illustrate how the categorical perception of speech depends not only on the physical (acoustic) properties of speech stimuli but on the knowledge of the listener as well. In other words, our purpose is to underline the fact that speech perception is a complex, multi-level, *interactive* process (McClelland and Elman 1986a, 1986b; Samuel 1990; Caplan 1992) in which *both* the so-called «lower-level variables» (=the phonetic properties of speech signals) and the «higher-level variables» (=the general knowledge background of listeners) interfere (see, e. g.: Fry 1970, 31, 47 and ff.; Studdert-Kennedy 1974, 2350; Pisoni and Sawusch 1975, 16-17). The latter variables are responsible for the fact that the same stimulus can be categorized in different ways, either by the same listener in different circumstances or even by different listeners as well. Examples of the higher-level variables are the ones taken into consideration in our study: LINGUISTIC EXPERIENCE and PHONETIC TRAINING of listeners.

The linguistic experience effect

The effect of linguistic experience on the categorical processing of speech has been in evidence since some «classical» experiments, such as Miyawaki, Strange, Verbrugge, Liberman, Jenkins and Fujimura's (1975), demonstrated that a listener can discriminate between two sounds (phonetically different) only if the acoustic differences reflect a phonemic contrast in his/her mother-tongue. If the two sounds do not correspond to a phonemic contrast – as is the case for the allophones of one phoneme in a given language –, a native listener of this language will not be able to discriminate between them; he/she will perceive them «as the same sound».

Recent developmental studies suggest that this intercategorical discrimination is the result of one's exposure to a particular phonological system during the first months of life (Kuhl 1993, 1995; Kuhl, Williams, Lacerda, Stevens and Lindblom 1992; Lacerda 1995; Lacerda and Sundberg 1997).

The linguistic experience effect has become an interesting topic of discussion and research, largely supported by experimental work, known as the «cross-linguistic speech perception» issue (Strange, ed. 1995).

The phonetic training effect

The interference of phonetic training on speech perception is assumed to have an opposite effect. Experimental studies, like those of Hillenbrand, Canter and Smith (1990), have suggested that listeners with phonetic training (professional linguists and phoneticians, for instance) are able to establish intraphonemic discrimination. This means that even if a phonetic contrast does not represent a phonemic contrast in a given language, a listener with phonetic

training will perceive both sounds of such a contrast as different (in the same way as a child would perceive them before his/her linguistic exposure determined his/her selective sensitivity to the functional contrasts of his/her mother-tongue only).

1.2. An allophonic contrast of Portuguese: the «voiced fricatized stops»

To illustrate these theoretical issues, we have chosen the case of an allophonic contrast in contemporary European Portuguese (henceforth, Portuguese): the so-called «voiced fricatized stops».

The consonantal system of Portuguese includes a series of three voiced stops (/b d g/), that can be phonetically realized either as «true» stops, with a silent interval during the total closure of the vocal tract (=[b d g]), or as a kind of «continuant stops» (=[β ð ɣ])[1]. The latter have been traditionally referred to as «voiced fricatized stops»[2] and are phonetically characterized by the lack of the silent interval: instead of it, due to the incomplete closure of the vocal tract during their articulation, they present a harmonic-like amount of acoustic energy of low intensity and transient configuration (Andrade, Andrade and Viana 1978; Viana 1984; Veloso 1995).

Since the silent interval is supposed to be one of the main acoustic cues to the identification of a stop (Ainsworth 1976, 76; Dorman, Raphael and Liberman 1976; Repp, Liberman, Eccardt and Pesetsky 1978; Kent and Read 1992, 110), its absence in these allophones leads us to the expectation that, if no other variables – such as linguistic knowledge, for example – interfere in the processing of these stimuli, they will be perceived as continuant consonants.

2. HYPOTHESES

Our hypotheses – based on the assumption that higher-level variables do interfere with our processing of speech stimuli and according to what has been said above – are the following:

(1) – native listeners of Portuguese will not discriminate the «fricatized stops» [β ð ɣ] from non-continuant [b d g] stops; non-native listeners will discriminate them;

(2) – phonetically trained Portuguese listeners will discriminate [β ð ɣ] from [b d g] more often that Portuguese listeners without phonetic training.

3. PROCEDURE

3.1. Stimuli

The stimuli for our study were made up of 12 nonwords extracted from real Portuguese speech. All nonwords had the invariant phonetic structure [aCá], where the intervocalic consonant could be one of the following: [p t k b d g v z ʒ β ð ɣ].

[1] Other allophonic possibilities of /b d g/ can be found in Portuguese; they will not be taken into consideration in this study (see Andrade, Andrade and Viana 1978; Viana 1984; Veloso 1995).

[2] This is a rough English translation of the most current and traditional Portuguese term *«oclusivas sonoras fricatizadas»;* this traditional designation will not be discussed here.

3.2. Subjects

The subjects were 17 adult individuals, divided into four groups, according to the combination of the two higher-level variables taken into consideration in this study:

- Group 1: native listeners of Portuguese without phonetic training (N=6);
- Group 2: native listeners of Portuguese with phonetic training (N=3);
- Group 3: non-native listeners of Portuguese without phonetic training (N=3);
- Group 4: non-native listeners of Portuguese with phonetic training(N=5);

The non-native listeners of Groups 3 and 4 were native speakers/listeners of languages for which descriptions of the «fricatized» stops have not been found: in Group 3 we found 2 native speakers/listeners of German and one of Italian; Group 4 included 3 native speakers/listeners of Swedish, 1 of German and 1 of Estonian.

3.3. Methodology

The stimuli were presented to all subjects through binaural headphones, in individual testing sessions that took place in a quiet room. Subjects of Groups 1, 2 and 3 were tested in individual sessions; subjects of Group 4 were tested in a simultaneous test session. All subjects were asked to identify the consonants they heard between the two vowels of each nonword. Subjects without phonetic training (Groups 1 and 3) were asked to make an orthographic transcription of the consonant, which was transformed into a phonetic transcription by the experimenter immediately after each session. Subjects with phonetic training (Groups 2 and 4) explicitly identified the manner of articulation of the consonants, answering a closed questionnaire where they were to assign one of two options («Stop» or «Continuant») to each consonant.

Each nonword was presented to each subject 3 times (12 nonwords X 3 presentations each = 36 stimuli); the order of the presentation of the stimuli was determined at random. Between one nonword and the following, a silent pause of 4 s was introduced. After each block of 6 stimuli, a longer silent pause of 12 s was introduced. In the questionnaire forms, each nonword corresponded to a separate answer line; answer lines were grouped in blocks of 6 lines each, in order to correspond to the order of stimuli and pauses presented through the headphones.

4. RESULTS

According to our hypotheses, we looked specifically at the answers that identified the realizations of /b d g/ as continuant consonants. The results are displayed in Table 1.

TABLE 1. Identification (%) of non-continuant and continuant («fricatized») stops as continuant consonants in the four groups of subjects

	[b]	[β]	[d]	[ð]	[g]	[ɣ]
GROUP 1	5,6	0	0	0	0	0
GROUP 2	33,3	33,3	22,2	66,7	0	0
GROUP 3	44,4	88,9	0	66,6	0	11,1
GROUP 4	6,7	73,3	0	100	0	20

Group 1: native listeners of Portuguese without phonetic training; Group 2: native listeners of Portuguese with phonetic training; Group 3: non-native listeners of Portuguese without phonetic training; Group 4: non-native listeners of Portuguese with phonetic training

Subjects with phonetic training (Groups 2 and 4) gave their answers as explicit identifications of the manner of articulation; subjects without phonetic training (Groups 1 and 3) presented orthographic transcriptions of the consonants.

As was predicted by our hypotheses, non-native listeners – either with (Group 4) or without (Group 3) phonetic training – show a general stronger tendency to identify the «fricatized» stops as continuants. This tendency is particularly clear in the cases of [β] and [ð], since [ɣ] seems to have a specific status in this matter (see discussion, below). Among the native listeners without phonetic training (Group 1), no answers identifying [β ð ɣ] as continuants were obtained.

If we compare the results of the identification of the manner of articulation of [β ð ɣ] among the subjects without phonetic training (Group 1 *vs* Group 3), we can see that the identification of [β] and [ð] as continuants is significantly much higher in the non-native listeners for each consonant: [β]: t=12,22; d.f.=7; p <0,0005; [ð]: t=5,292; d.f.=7; p <0,005. The difference found in the identification of the manner of articulation of the velar «fricatized» stop between these two groups of subjects did not present the same significance: [ɣ]: t=1,528; d.f.=7; p <0,10.

Looking at the differences in the identification of the manner of articulation of [β] and [ð] between native subjects with and without phonetic training (Group 1 *vs* Group 2), even if one can recognize a slightly higher tendency among the phonetically trained subjects to identify the «fricatized» stops as continuants, only one significant difference was found, affecting the consonant [ð]: [β]: t=1,528; d.f.=7; p <0,10; [ð]: t=3,055; d.f.=7; p <0,01.

5. GENERAL DISCUSSION AND CONCLUDING REMARKS

Any conclusions extracted from this study are restricted to the limited number of subjects and stimuli used in our perceptual tests and need a further, deeper reevaluation with a larger number of subjects and more powerful statistical analysis. Therefore, any definitive conclusions will be avoided here.

However, some final remarks can be drawn based on the limited data we have obtained and on the rough analysis to which they were submitted.

At first sight, our results seem to support the previous expectations and hypotheses: on the one hand, significant differences in the identification of the manner of articulation of the «fricatized» stops were found between non-native and native listeners, in the sense that non-native listeners identified these consonants as continuants more often than native listeners did; on the other hand, phonetically trained subjects have been shown to be more proficient in the intraphonemic distinctions between non-continuant and continuant realizations of voiced stops.

As far as differences between non-native and native listeners are concerned, the above mentioned remarks relate to the case of listeners without phonetic training only (Groups 1 and 3). These listeners represent the most common situation of speech perception, i. e., the «real speech communication situation» (when speech is transmitted from one speaker to another in order to convey some linguistic message, it normally takes place between «naïve» subjects, who generally do not have any explicit phonetic knowledge). If we look at the listeners with phonetic training (Groups 2 and 4), and we compare non-native with native listeners, no significant differences were found in the identification of the manner of articulation of [β ð ɣ] between these two groups of subjects. This discrepancy between the two comparisons (non-native without phonetic training *vs* native without phonetic training AND non-native with phonetic training *vs* native with phonetic training) could be attributed to the specific kind of knowledge that seems to be activated in the task of explicit identification of a phonetic property like manner of articulation. As was said earlier, phonetically trained subjects were asked for an explicit identification of the manner of articulation, whereas subjects without phonetic training were asked for an orthographic transcription. These are two different skills that implicate different kinds of knowledge. One of the major differences between the two is that an orthographic transcription always forces the transcriber to assign a stimulus to a category, independently of some individual variation among stimuli, and any comparison of results obtained by means of these two different tasks must have this discrepancy in mind.

Thus, going back to our previous hypothesis that phonetically trained subjects would identify [β ð ɣ] as continuants more often than subjects without phonetic training, we were able to observe a somewhat higher tendency towards this effect among the phonetically trained listeners. Anyway, this tendency was not shown to be significant by the statistical analysis to which our data were submitted. This lack of significance suggests the necessity of a deeper analysis of our results and the repetition of our tests with other, larger groups of subjects. Only this reevaluation will be able to shed more light on the relation between the two higher-level variables that are taken into consideration in this study.

All these concluding remarks find some application as far as the processing of [β] and [ð] is concerned. The velar stop /g/, either realized as [g] or as [V], is shown to be processed in a very special manner. In fact, whatever subjects processed its allophones and whatever allophones were identified ([g] or [V]), the velar consonant was almost always identified as a non-continuant stop (with the rare exceptions of 20% of the answers of the non-native listeners with phonetic training, all of them given by the same single subject, and the 11,1% of the answers of the non-native listeners without phonetic training, representing one single answer of one subject only). This stability in the processing of this consonant is probably due to some phonetic properties perhaps related to its place of articulation, and deserves a further phonetic exploration as an independent topic.

Finally, as a main concluding remark – always bearing in mind the above mentioned limitations of this work –, we would like to say that this study can be added to those that bring some supportive arguments to the view that speech perception is, in fact, a complex, multi-level, interactive process which comprehends a lot of variables, encoded not only in the acoustic signal of speech, but in the knowledge and the multiple skills of the listener as well.

ACKNOWLEDGMENTS

This paper is based on a master-equivalent thesis whose supervisor was Professor Maria da Graça Pinto

(University of Porto). I thank her for all her support during the research which led to this dissertation, as well as for her useful comments on an earlier draft of this paper. A substantial part of the experimental work which is reported here was carried out at the Phonetics Laboratory of Stockholm University, thanks to a scholarship from The Swedish Institute (1992) and the cooperation of Professor Francisco Lacerda (Stockholm University). Many thanks are also due to Professor Belinda Maia (University of Porto) for her kind help with the final English form of this text.

All these acknowledgments notwithstanding, all faults of this work are, of course, exclusively my own.

REFERENCES

Ainsworth, W. A. 1976. *Mechanisms of Speech Recognition.* Oxford, Pergamon Press.

Andrade, A.; Andrade, E.; Viana, M.C. 1978. *A fricatização das oclusivas sonoras em português.* Lisboa, INIC/Centro de Linguística da Universidade de Lisboa. Relatório do Grupo de Fonética e Fonologia, nº 3.

Caplan, D. 1992. *Language – Structure, Processing, and Disorders.* Cambridge (Mass.), The MIT Press.

Dorman, M. F.; Raphael, L. J.; Liberman, A. M. 1976. Further Observations on the Role of Silence in the Perception of Stop Consonants. *Speech Research: Haskins Laboratories Status Report, SR-48,* 199-207.

Fry, D. B. 1970. Speech Perception and Reception. In J. Lyons ed. *New Horizons in Linguistics.* London, Penguin Books, 29-52.

Hillenbrand, J.; Canter, G. J.; Smith, B. L. 1990. Perception of intraphonemic differences by phoneticians, musicians, and inexperienced listeners. *Journal of the Acoustical Society of America, 88(2),* 655-662.

Kent, R. D.; Read, C. 1992. *The Acoustic Analysis of Speech.* San Diego, Singular.

Kuhl, P. K. 1993. Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics, 21(1-2),* 125-139.

Kuhl, P. K. 1995. Mechanisms of developmental change in speech and language. In *Proceedings of the XIII.th International Congress of Phonetic Sciences,* ed. by K. Elenius and P. Branderud, Stockholm, Stockholm University/KTH, Vol. 2, 132-139.

Kuhl, P. K.; Williams, K. A.; Lacerda, F.; Stevens, K. N.; Lindblom, B. 1992. Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science, 255,* 606-608.

Lacerda, F. 1995. The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In *Proceedings of the XIII.th International Congress of Phonetic Sciences,* ed. by K. Elenius and P. Branderud, Stockholm, Stockholm University/KTH, Vol. 2, 140-147.

Lacerda, F.; Sundberg, U. 1997. The phonetic basis of language acquisition. Paper presented to the 5.th International Congress of the International Society of Applied Psycholinguistics (Porto, 25-27 June, 1997).

McClelland, J. L.; Elman, J. L. 1986a. Interactive Processes in Speech Perception: The TRACE Model. In McClelland, J. L.; Rumelhart, D. E.; The PDP Research Group. *Parallel Distributed Processing – Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models.* Cambridge (Mass.), The MIT Press/Bradford Books, 58-121.

McClelland, J. L.; Elman, J. L. 1986b. The TRACE Model of Speech Perception. *Cognitive Psychology, 18,* 1-86.

Miyawaki, K.; Strange, W.; Verbrugge, R.; Liberman, A.M.; Jenkins, J. J.; Fujimura, O. 1975. An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics, 18(5),* 331-340.

Pisoni, D. B.; Sawusch, J. R. 1975. Some Stages of Processing in Speech Perception. In A. Cohen, S. G. Nooteboom eds.*Structure and Process in Speech Perception.* Berlin, Springer-Verlag, 16-35.

Repp, B. H.; Liberman, A. M.; Eccardt, T.; Pesetsky, D. 1978. Perceptual Integration of Acoustic Cues for Stop, Fricative, and Affricate Manner. *Journal of Experimental Psychology: Human Perception and Performance, 4(4),* 621-637.

Samuel, A. G. 1990. Using Perceptual-Restoration Effects to Explore the Architecture of Perception. In G. T. M. Altmann ed. *Cognitive Models of Speech Processing – Psycholinguistic and Computational Perspectives,* Cambridge (Mass.), The MIT Press/Bradford Books, 295-314.

Strange, W. ed. 1995. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* Timonium (MD), York Press.

Studdert-Kennedy, M. 1974. The perception of speech. In T. A. SEBEOK ed. *Current Trends in Linguistics – Volume 12(1). Linguistic and Adjacent Arts and Sciences.* The Hague/Paris, Mouton, 2349-2385.

Veloso, J. 1995. *Aspectos da Percepção das «Oclusivas Fricatizadas» do Português.* Porto, Universidade do Porto – Faculdade de Letras. Master-equivalent thesis.

Viana, M.C. 1984. *Etude de Deux Aspects du Consonantisme du Portugais: Fricatisation et Dévoisement.* Strasbourg, Université des Sciences Humaines de Strasbourg. Doctoral dissertation.