

EFFECTS OF LANGUAGE CHANGE ON VOICE QUALITY. AN EXPERIMENTAL
CONTRIBUTION TO THE STUDY OF THE CATALAN-CASTILIAN CASE

Bernard HARMEGNIES (+), Marielle BRUYNINCKX (+), Joaquim LLISTERI (*), Dolores POCH (*)

(+)
(*)

Université de Mons, Belgium
Universitat Autònoma de Barcelona, Spain

ABSTRACT

Various experiments have dealt with the idea that voice quality could be influenced by the speaker's language. Most authors in this field have tried to objectivate voice quality by means of long term average spectra (LTAS). It has been shown [1] that their results - although apparently contradictory - reveal that part of the LTAS variability can be explained by language variations. Recently, a powerful method allowing an accurate evaluation of those effects has been developed [2] and applied to Dutch and French. The purpose of the present work is to replicate this study with a new pair of languages : Catalan and Spanish (Castilian). 10 bilingual speakers have produced, each, 4 readings of 2 (1 Castilian and 1 Catalan) balanced texts. A 400-line LTAS was computed for each utterance. The LTAS are compared by means of the SDDD dissimilarity index. The results confirm previous findings, i.e., the inter-language variability is greater than the intra-language variabilities.

1. INTRODUCTION

Several authors have suggested that the language spoken might influence the long term strain of speech. This idea has been mainly promoted by pedagogues interested in teaching languages. For instance, Tomatis [3] has introduced the idea that what he called the "envelope curves of languages" are determinant for the perception of any language, and should therefore be taken into account each time a student begins learning a new language. Some researchers interested in the study of speech signal have also carried out experiments aimed at describing the spectral long term features of several languages. However, their findings have turned out to be rather contradictory, since some of them reported significant language effects [4,5,6], although others did not [7,8]. In the field of speaker recognition, two separate sets of experiments dealing with various languages have led to apparently contradictory conclusions : in this respect, Nolan [9] emphasized the indirect conflict between Tosi's results [10] on the one hand and Majewski and Hollien's findings [11] on the other.

Harmegnies and Landercy [12] have criticized the descriptive research in this field, showing its lack of efficient ways of quantifying the observed variabilities. They have also suggested that, in some cases, at least a part of the

assumed language effects reported could in fact simply have been produced by interindividual variability. They have built, on this basis, a mixed quantitative/qualitative methodology which allowed them to carry out careful examinations of the differences between Dutch and French. Their results [1,2,12] confirmed the existence of a significant language-related long term spectral variability, but also showed that language effects are far smaller than individual effects.

Obviously, these findings are of some interest, since they make it possible to understand how various authors can have come to apparently contradictory conclusions. Nevertheless, they are based upon two languages only, i.e., Dutch and French. It may therefore seem highly desirable to apply the methodology to other languages.

This is precisely the purpose of this paper, in which we will try to study to what extent long term average speech spectra (LTAS) can be influenced by the use of Castilian (Spanish) as opposed to Catalan.

2. BILINGUALISM IN CATALONIA

Catalonia is a bilingual community where both Catalan and Castilian are used. Native Catalan speakers know Castilian and are able to use it with different degrees of fluency and accuracy. Native Castilian speakers living in Catalonia usually understand Catalan, but some of them do not speak it. There is therefore a great amount of sociolinguistic variability, ranging from quasi-monolingual Castilian to completely bilingual Catalan speakers.

However, due to the fact that Castilian is widely used by the media and is also taught at schools, it is almost impossible to find monolingual Catalan subjects.

Little work has been carried out on the analysis of phonetic performance of bilingual speakers. Experimental studies until now have dealt with the vocalic system [13] and with the formant pattern of lateral consonants [14,15]. They are mainly concerned with some of the subphonemic features related to Catalan "accent" when Catalan subjects use Castilian. The results show a high degree of influence of the first language over the second one.

3. EXPERIMENT

3.1. Subjects selection

10 bilingual subjects were chosen on the basis of a written questionnaire assessing their use of spoken and written Catalan and Castilian in all situations. They were chosen among students from language departments of the Universitat Autònoma de Barcelona. They were between 18 and 22 years old.

We were looking for subjects with a balanced use of both languages. Their recordings were carefully checked in order to ensure that they did not show strong accent features either in Catalan or in Castilian.

3.2. Corpus

Two phonetically balanced corpuses were produced: one for Castilian and one for Catalan. They lasted for about 20 seconds, read at the subject's natural rate.

3.3. Experimental setting

Each subject selected was asked to utter each text four times. The recordings took place in a sound-proof room. The subjects were sitting in front of the microphone, placed at a constant 20 cm from their lips. All texts were recorded on a REVOX A 77 recorder, by means of a SENNHEISER, MD 441 N cardioid microphone.

3.4. Acoustical analysis

The recordings were afterwards analysed by means of a Bruel Kjaer 2033 400 - channel FFT analyser. Its sampling frequency was set at 12.8 kHz, determining a constant 12.5 Hz resolution across the whole DC-5 kHz range of analysis. A linear averaging process was moreover used in order to compute one LTAS for each utterance.

The 80 (10 subjects x 4 utterances x 2 languages) so-obtained LTAS were then transmitted from the analyser to an IBM PC XT computer via a GPIB interface card, for further computations.

3.5. Comparison procedure

For the purpose of comparing possible effects of languages in our bilingual subjects, we performed, for each subject, 3 kinds of intra-speaker comparisons of the stored LTAS. Intra- as well as inter-language comparisons were performed. The first ones consisted in comparing Castilian LTAS with Castilian LTAS (CAST/CAST), and Catalan LTAS with Catalan LTAS (CAT/CAT), although inter-language comparisons consisted in comparing Catalan LTAS with Castilian LTAS (CAT/CAST). In the case of the intra language comparisons (CAT/CAT, as well as CAST/CAST), for each one of the 10 speakers, one comparison was performed for each possible non-redundant pair of his 4 LTAS in each language, i.e., 6 comparisons by subject and by language, thus 120 intra-speaker comparisons for the whole sample (10 subjects x 6 comparisons x 2 languages). In the case of the inter-language comparisons (CAT/CAST), each Catalan LTAS of each subject was compared to each Castilian LTAS of the same subject. This procedure

therefore resulted in 160 inter-language/intra-speaker comparisons (16 comparisons x 10 subjects).

Each individual comparison of a given LTAS with another was performed by means of the SDDD index [16], which measures the dissimilarity of the spectra under comparison: high SDDD values indicate considerable dissimilarities between spectra, and conversely. SDDD is defined as:

$$SDDD_{SS'} = \left(\frac{1}{K} \sum_{i=1}^K (S_i - S'_i - M_d)^2 \right)^{1/2} \quad (1)$$

where S and S' are the spectra being compared, each one defined by its s_i and s'_i level values for each of the k frequency components, and M_d is the mean for the $S_i - S'_i$ differences. SDDD is insensitive to variations in the overall levels of the compared spectra, which, therefore, do not need any intensity-normalization prior to the computations.

4. RESULTS

Tables 1, 2 and 3 present a summary of the observed statistical distributions for each kind of comparison (respectively, CAT/CAT, CAST/CAST, and CAST/CAT). It can be seen from tables 1 and 2 that the SDDD values drawn from intra languages comparisons appear to be within the same range as values previously reported by Harmegnies [1], on the occasion of intra-language comparisons of various French LTAS. The average SDDD values he observed scored between 2.20 and 2.66, depending on the corpus content, and our present data show average SDDD respectively equal to 2.47 (CAT/CAT) and 2.53 (CAST/CAST). On the other hand, the average SDDD value drawn from our inter-language comparisons is higher (2.94). Another striking feature of our data is the considerable amount of inter-speaker variability. The individual average SDDD values vary from 1.92 up to 3.39 (CAT/CAT), 2.02 up to 3.54 (CAST/CAST), and 2.26 up to 3.75 (CAT/CAST). Nevertheless, it seems that, for a given subject, a high SDDD value in a comparison type is associated with high SDDD values in the others, and conversely.

Speaker	Low. value	High. value	Mean	St. dev.
1	1.790	2.464	2.131	.272
2	1.921	2.350	2.210	.160
3	1.829	2.455	2.104	.250
4	2.026	2.451	2.172	.138
5	1.751	2.441	2.185	.286
6	1.962	2.648	2.201	.226
7	1.720	2.569	2.024	.275
8	2.488	4.412	3.355	.727
9	2.770	3.927	3.329	.365
10	2.465	4.139	3.543	.566
Low.value	1.720	2.350	2.024	.138
High.value	2.770	4.412	3.543	.727
Mean			2.525	.327
St. dev.			.583	.175

Table 1: Lower values, higher values, means and standard deviations of the SDDD values drawn from intra languages CAST/CAST comparisons

Speaker	Low. value	High. value	Mean	St. dev.
1	1.659	2.242	1.940	.175
2	1.922	2.866	2.316	.340
3	1.948	2.604	2.276	.239
4	2.633	3.389	2.963	.238
5	1.873	2.077	1.968	.074
6	1.876	2.500	2.097	.209
7	1.664	2.149	1.917	.148
8	2.170	4.714	3.390	.954
9	2.453	3.058	2.748	.210
10	2.576	3.485	3.054	.384

Low.value	1.659	2.077	1.917	.074
High.value	2.633	4.714	3.390	.954
Mean			2.467	.297
St. dev.			.505	.235

Table 2 : Lower values, higher values, means and standard deviations of the SDDD values drawn from the intra-language CAT/CAT comparisons.

Speaker	Low. value	High. value	Mean	St. dev.
1	1.913	2.647	2.263	.171
2	2.146	3.051	2.633	.257
3	2.278	3.904	3.014	.473
4	2.349	3.176	2.811	.218
5	2.466	3.152	2.853	.198
6	2.277	2.763	2.536	.130
7	2.250	3.162	2.760	.234
8	2.307	4.695	3.286	.774
9	3.204	4.497	3.753	.439
10	2.167	4.603	3.485	.703

Low.value	1.913	2.647	2.263	.130
High.value	3.204	4.695	3.753	.774
Mean			2.940	.360
St. dev.			.431	.217

Table 3 : Lower values, Higher values, means and standard values of the SDDD values drawn from the inter-languages CAST/CAT comparisons.

In order to evaluate the validity of these preliminary observations, we performed two specific statistical analyses.

The first one was a two-way analysis of variance aiming at studying the structures of variability among the whole set of SDDD values. The 3 comparison types were considered as the 3 levels of a fixed dimension, although the 10 subjects were considered as 10 levels of a random dimension. The analysis revealed significant comparison type-, subject-, and interaction effects (respectively : $F=12.28, \alpha < .001$; $F=34.20, \alpha < .001$; $F=3.07, \alpha < .001$).

We had the second analysis differ from the first one only by suppressing the data set drawn from inter-language comparisons. We obtained, this way, a balanced two-way, mixed design, with the subjects as a random dimension, and two intra language comparison types (CAT/CAT and CAST/CAST) as the two levels of a fixed dimension. This analysis confirmed strong subject- and interaction effects (respectively : $F=20.84, \alpha < .001$; $F=2.63, \alpha = .009$), but did not reveal any effects relating to the comparison types ($F=.21, \alpha = .642$).

The first analysis of variance exhibits a strong inter-comparison-conditions effect. The second analysis of variance, dealing with intra-language comparisons only, does not

reveal any such effect. One may thus conclude that the effects observed on the occasion of the first analysis are due only to the presence of the inter-language data in the analysis design. A significant language effect can therefore be inferred : the SDDD values drawn from inter-language comparisons score higher than those drawn from intra-language (CAT/CAT and CAST/CAST) comparisons.

Nevertheless, both analyses lead to the common conclusion that a great amount of variability derives from the subjects themselves. Moreover, the significant interaction effect shows that the influence exerted by languages upon LTAS depends on the subjects. Interpreted in terms of voice consistency, these data suggest that the speakers could be characterized by their own inter-language coherence.

5. CONCLUSION

The conclusions of the present research are in good agreement with those of the others previously carried out at the University of Mons [1,2,12]. They confirm that the LTAS is sensitive to language changes. Moreover, they can be considered as stronger, since they are derived from experimental conditions which are likely to minimize the inter language LTAS variability : on the one hand, as already observed [1], the use of bilingual subjects may have resulted in diminishing the between - languages differences; on the other, one may expect that, as both languages under investigation belong to the same Romance group, they are very similar in qualities.

It is also to be noticed that the language effect has been revealed in spite of an important variability relating to the subjects themselves. Inter subject variabilities of this kind have already been pointed out [1,2]. It has been suggested elsewhere [17] that, as a general rule, speakers could be characterized by their own degree of voice coherence. The strong interaction- and between speaker effects revealed by our experiment suggest an extension of the concept, leading to the one of inter language coherence.

If the inter language coherence varies (as our results suggest) one should, finally, ask why. A possible explanation could be that various degrees of bilingualism result in various degrees in inter language coherence. Therefore, we think that further studies should reach a more refined assessment of the degree of bilingualism exhibited by the subjects. This could lead, e.g., to investigations involving groups showing different abilities in Catalan and Castilian.

It is our hope that this kind of research finally contributes to a better understanding of the phonetic processes underlying bilinguals' use of languages.

6. REFERENCES

- [1] HARMEGNIES, B., Contribution à la caractérisation de la qualité vocale; analyses plurielles de spectres moyens à long terme de parole, Dissertation doctorale, Université de Mons, 1988.

- [2] HARMEGNIES, B., LANDERCY, A., BRUYNINCKX, M., "An experiment in inter-languages speaker recognition using the SDDD index", Proceedings of the 11th International Congress of the Phonetic Sciences, Tallinn, 241-244, 1987.
- [3] TOMATIS, L.A., L'intégration des langues vivantes, Paris, Soditap, 1970.
- [4] TARNOCZY, T., "Détermination du spectre de la parole avec une méthode nouvelle", Acustica, 8, 392-395, 1958.
- [5] TARNOCZY, T., "Sprachliche vershieden erhalten durch sprechchormethode", 4th Int. Congr. on Acoustics, Copenhagen, 627, 1-4, 1962.
- [6] BANUS-TEROL, V., "Weighted average spectrum of human speech : an approach", Proc. of the 7th Int. Congr. on Acoustics, 253-256, 1971.
- [7] LORAND, P., MINIER, G., SALIOU, A., "Détermination du spectre à long terme de la parole (langue française)", Recherches en Acoustique, II, CNET, 15-25, 1975.
- [8] BOULLOSA, R.R., PEREZ RUIZ, S.J., "Long-term average spectrum of Spanish", Appl. Acoust., 17, 3, 1984.
- [9] NOLAN, F., The Phonetic bases of speaker recognition, Cambridge, Cambridge University Press, 1983.
- [10] TOSI, O., Voice identification : Theory and Legal Applications, Baltimore, University Park Press, 1979.
- [11] MAJEWSKI, W., HOLLIEN, H., "Euclidean distances between long term speech spectra as a criterion for speaker identification", In FANT, G., (ed), Speech Communication, 3, Stockholm, Almqvist and Wiksell, 303-310, 1975.
- [12] HARMEGNIES, B., LANDERCY, A., "Language features in the long-term average spectrum", Revue de Phonétique Appliquée, 73-5, 69-80, 1985.
- [13] LLISTERRI, J., POCH, D., BALARI, S., "Estructuraciyn fonotica de la segunda lengua en hablantes bilingües", Estudios de Lingüística, Universidad de Alicante, forthcoming.
- [14] COMAS, T., Caracteritzaciyn acústica de les realitzacions de /l/ en parlants bilingües català-castellà, Unpublished thesis, Universitat Autònoma de Barcelona, 1986.
- [15] MARTINEZ, G., "Caracterizaciyn acústica de las realizaciones en hablo continua de la consonante lateral /l/ en hablantes bilingües", XVIII Simposio de la Sociedad Española de Lingüística, Madrid, 1988.
- [16] HARMEGNIES, B., SDDD a new dissimilarity index for the comparison of speech spectra, Pattern recognition letters, 8, 153-158, 1988.
- [17] HARMEGNIES, B., LANDERCY, A., "Intra-speaker variability of the Long Term Speech Spectrum", Speech Communication, 7, 81-86, 1988.