Vowel formant structure and speaker identification. A perceptual study

Mª José Albalá[4], Elena Battaner[3], Juana Gil[4], Joaquim Llisterri[1], María Machuca[1], Victoria Marrero[2], Carme de la Mota[1], Montserrat Marquina[1], Antonio Ríos[1]

[1] Universidad Autónoma de Barcelona [2] Universidad Nacional de Educación a Distancia [3] Universidad Rey Juan Carlos [4] Consejo Superior de Investigaciones Científicas

This study belongs to VILE Project (Acoustic and Perceptual Study of Variation within and between Speakers in Spanish[1]), whose goal is to validate -acoustically and perceptually- relevant cues to establish the individuality of speakers, in order to apply the results to automatic speech recognition and to judicial practice (Battaner et al. 2003; Marrero et al. 2003). On this occasion we'll present the results of a natural vowel 2IAX discrimination test. Stimuli were extracted from the acoustic database used in VILE-1st part. 40 pairs (4 vowels -/i,e,a,o/- * 2 accents –stressed/unstressed- * 5 acoustic parameters -F0/F1/F2/F3/F4-), as similar as possible in all acoustic parameters except for the one under study were presented to 46 judges (34 women and 12 men) aged 18-35. Stimuli were delivered via headphones in soundproof radio cabin using *Paradigm* software (Lopez-Bascuas et al. 1999). Results show 94.53% of hits when stimuli were the same, and 72,97% for different stimuli. Concerning the four vowels, the proportion of errors was 28.47 for /o/; 27.01 /a/; 26,48% /e/; 26,14 /i/. As for the acoustic parameters, the biggest error rate corresponds to F4 (39.97%), F3 (35.35%) and F1 (31.09). On the contrary, F0 ratio is only 13.73% and F2 14.70%. So, our findings seem to indicate the prevalence of F0 and F2 as main cues for speaker identification across different vowel categories. In the conclusion, implications of these results for future research are discussed.