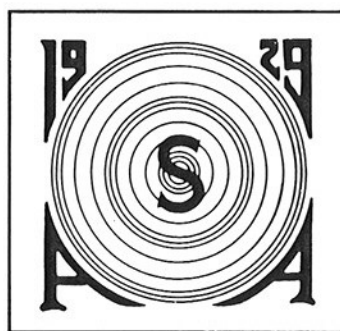


THE JOURNAL of the Acoustical Society of America

Vol. 95, No. 5, Pt. 2, May 1994



**Program of the
127th Meeting**



**Massachusetts Institute of Technology
Cambridge, Massachusetts
6-10 June 1994**

Table of Contents on p. v

Published by the Acoustical Society of America through the American Institute of Physics

the vectors of each speaker through the general SOM prototype. Every individual SOM is fine tuned by means of the LVQ3 algorithm. In the verification phase, a short parametrized speech interval of 3-s duration from the speaker under verification is compared to both the general prototype and the speaker's individual SOM. So two accumulated distortions are calculated. The decision to accept or reject a speaker depends on whether the speaker's voice matches to the claimed speaker's individual SOM or to the general prototype SOM. The MFCC, LPCC, and LSP coefficients were used in these studies as distinctive features. Personal SOMs were trained using 100 s of speech from each speaker. A conventional LBG algorithm was also applied. LVQ3 initialized by the SOM codebook was always more efficient than LVQ3 initialized by the LBG codebook using all three kinds of above-mentioned features.

2pSP23. Speaker sex and formant frequencies of American [ɹ]. Robert Hagiwara (Phon. Lab., UCLA Dept. of Linguist., 405 Hilgard Ave., Los Angeles, CA 90024)

Phonetic studies of speaker sex have primarily focused on phonation and vowel formants—very little is known about the effect of speaker sex on consonants and consonant-like sounds. Similarly, studies of American [ɹ] have not explored whether and how speaker sex affects the acoustic character of this sound. As part of a larger study of speaker sex and approximant [ɹ], 10 speakers (5 male, 5 female) of American English read three repetitions of 69 target words presented randomly in a frame. In addition, gross tongue shape used for the various allophones of [ɹ] was determined for each speaker by inserting a probe into the mouth during articulation. At least three tongue shapes can be discriminated using this method. In this presentation, the formant frequency data for the vowel tokens, including syllabic [ɹ], are examined in light of the following issues: the location of syllabic [ɹ] in a traditional vowel space ($F_1 \times F_2$); the quantitative effect of speaker sex on the frequency of the first three formants of syllabic [ɹ]; and whether there is any sex bias in the distribution of tongue shapes used to produce this sound, and what effect this has on its acoustic quality.

2pSP24. Gemination of Italian stops. Riccardo Rossetti (INFOCOM Dept., Univ. of Rome "La Sapienza," via Eudossiana 18, 00184 Rome, Italy)

An acoustic study of geminated and nongeminated Italian stops was carried out. The analysis was performed on 648 VCV syllables, where V is one of the vowels [i, a, u] and C is one of the consonants [p, t, k, b, d, g] (in geminated and nongeminated form). The syllables were uttered by 6 speakers, 3 repetitions. The analysis of the segments durations corresponding to the different articulatory phases of the VCV structure showed a strong correlation between the presence of gemination and the first vowel and occlusive silence durations. Statistical analyses using an *a posteriori* classifier yielded to a classification score of 96.1%. An analysis of the properties of the burst (energy, power, and spectral shape) showed that there is no correlation between these and the presence of gemination. An automatic software system, which allows the modification of a single/geminated VCV utterance into a geminated/single one, was developed. Informal perceptual tests, in which the stimuli were obtained by using the previous program, validated the significance of the first vowel and the occlusion durations as acoustic correlates of gemination.

2pSP25. Acoustic correlates of distinctive features of Italian stops. Armando Vannucci (INFOCOM Dept., Univ. of Rome "La Sapienza," via Eudossiana 18, 00184 Rome, Italy)

The object of this study was the determination of acoustic correlates of two distinctive features of Italian stops related to voicing and point of articulation. The analysis was carried out on 648 VCV syllables. The correlates found for voicing were the sound bar during the voiced stops occlusion and the low-frequency (LF) burst energy. Informal perceptual tests showed the relevance of only the second correlate. As to the point of articulation, the attention was focused both on formants transition from the first vowel to the following stop and on the spectral shape of the burst. For the formants transition, different groups of templates were obtained for the different vowels. The burst spectrum was then analyzed to obtain a gross representation through four cen-

troids defining three segments with different slopes. *A posteriori* classification based on either voicing (sound bar and LF burst energy) or point of articulation (gross spectral shape) yielded on the previous database to a classification score of 88% and 87%, respectively. These strategies were implemented in a system used for the recognition of Italian stops in VCV structures. An inter-linguistic study aimed at comparing Italian dental stops with British English corresponding alveolars in progress.

2pSP26. Influence of the vocal effort on vowels. Maria-Gabriella Di Benedetto (Dept. INFOCOM, Università degli Studi di Roma "La Sapienza," via Eudossiana 18, 00184 Rome, Italy) and Jean-Sylvain Liénard (LIMSI-CNRS, 91403 Orsay Cedex, France)

Vocal effort is one of the less studied sources of speech variability. An investigation of its effects on French vowels was carried out. Twelve French vowels uttered in isolation by 13 speakers, 7 females and 6 males, were recorded. The subject was instructed to repeat the vowel pronounced by the operator with three degrees of vocal effort induced by the distance between the speaker and the operator (30 cm and 1.5 and 7 m); the variations kept within the interval encountered in the usual living conditions, far from the whispered or shouted types of voice. Some of the speakers were recorded twice, 6 months apart. The resulting 720 tokens were assessed by a separate group of listeners who judged vowel identity, perceived speaker's gender, and perceived vocal effort. Results of acoustic analysis show a change in the global slope of the spectrum and a significant increase of the F_1 frequency correlated with the degree of vocal effort, with some dependency on the particular vowel and speaker. The average increase is about 50 Hz between the two extreme speaking conditions, corresponding to an interval of about 12 dB in the overall level. [Work supported by a CNRS-CNR cooperation.]

2pSP27. Abstract withdrawn.

2pSP28. Trans-syllabic spread of consonantal effects on vowel formant frequencies. Sarah Hawkins and Andrew Slater (Dept. of Linguist., Univ. of Cambridge, Sidgwick Ave., Cambridge CB3 9DA, UK)

Formant-based synthetic speech is less robust in noise than natural speech and is often criticized as "robot-like." One reason may be the failure to model nonlocal spectral variation induced by segments in other syllables. This study investigates how a single consonant can affect vowel formant frequencies in nonadjacent as well as adjacent syllables. Sequences of /ə u i ə/ with intervening consonants were embedded in carrier phrases to give quasimeaningful sentences. The medial consonant was /z/ or /r/. Flanking consonants were either all /b/, or /d/ in the syllables before and after the medial consonant (/bəbu(?)ibəb/, /bədu(?)idəb/). Primary stress was on either the second syllable, or the first and third. Steady-state or mid-point frequencies of F_1 – F_3 were