Studi italiani di linguistica teorica ed applicata, anno XVII, 1988, numero 2-3

MARIA-GABRIELLA DI BENEDETTO *Roma*

THE INFLUENCE OF THE FIRST FORMANT FREQUENCY TRAJECTORY SHAPE ON THE PERCEPTION OF THE TENSE/LAX DISTINCTION OF AMERICAN ENGLISH AND ITALIAN VOWELS

1. INTRODUCTION

The feature [tense] has been broadly used in models of generative phonology to indicate the linguistic distinction between some consonants (for example [t] vs [d]) and also some vowels (for example [e] vs [ɛ] in Italian, as can be found in Muljačič (1972), Brozovič (1967), Saltarelli (1970), and Tonelli (1984)). Various attempts have been made to correlate this distinctive feature to some specific property. From an articulatory point of view, the question is controversial. One hypothesis is the major tenseness of some muscles of the vocal tract during the production of the sounds defined as tense. For example, the tension of the tongue against the hard palate may seem one property related with the production of the [t,d] distinction, but there is no clear evidence of an effectively well-defined muscular distinction in the production of tense and lax vowels. From an acoustic point of view, properties based on measurements of the time-varying spectrum may be related to the feature [tense], namely some properties of the formant trajectories (Lehiste and Peterson, 1961) and the variation of the relative distance between some of the formants, with respect to the central acoustic position of the vowel schwa. It appears then that the feature [tense] may be characterized by different acoustic and articulatory manifestations. As proposed by Stevens et al. (1986), this makes this feature a possible candidate for being called a "cover feature" in distinctive feature theory. From a perceptual point of view, the variation of some specific property of synthetic stimuli could lead to a perceptual distinction between tense and lax sounds (Huang, 1985), after the spectrum has been processed and normalized by the auditory system. Furthermore, the evolution of the languages shows that different languages have made different use of the tense/lax distinction, so that sounds effectively linguistically distinct along the tense/lax dimension have been created or progressively deleted. For example, in Italian, minimal pairs based on the difference between two vowels of a tense/lax pair ([e] vs [ɛ], in the words [peska], 'fishing', vs [pɛska], 'peach') have disappeared in most dialects and the distinction is rarely made in natural speech.

In this paper, perceptual experiments carried out using synthetic CVC syllables in order to specify perceptual correlates of the tense/lax distinction will be described. Two languages, American English and Italian, characterized by two different vowel systems and by a different phonological use of the tense/lax distinction will be considered, and the results obtained with subjects native speakers of these languages will be compared. Note that the Italian vowel system includes two pairs [e,ɛ] and [o,ɔ] which are described phonologically as tense/lax pairs. Figure 1 shows an example

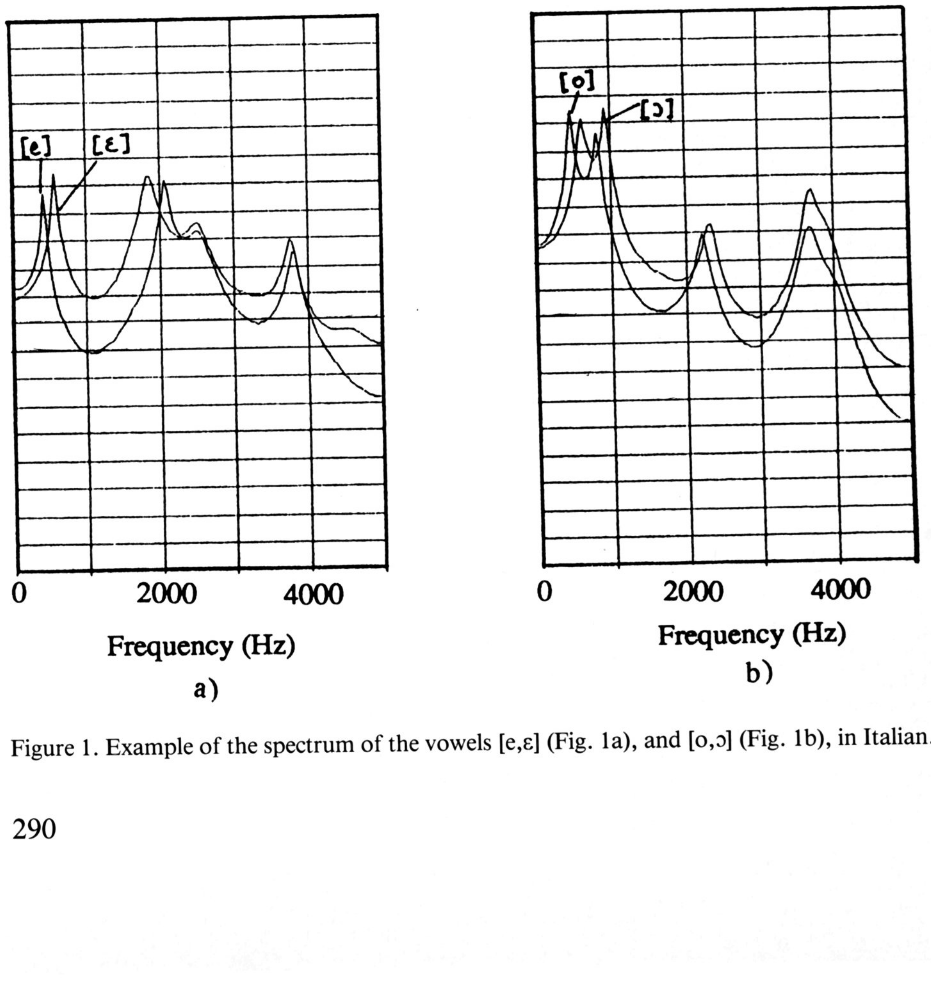

Figure 1. Example of the spectrum of the vowels [e,ɛ] (Fig. 1a), and [o,ɔ] (Fig. 1b), in Italian.

of a linear predictive coding (LPC) spectrum of these pairs for the same speaker ([e,ɛ] in Fig. 1a and [o,ɔ] in Fig. 1b), obtained by using 16 LPC coefficients.

2. Description of the experiment

The aim of the perceptual experiment described in the present paper was to investigate the influence of the first formant frequency trajectory shape on the perception of the tense/lax distinction of American English and Italian vowels. This investigation was part of a larger study on the effect of F1 trajectory shape modifications on vowel identification (Di Benedetto, 1987). The stimuli used were all dVd synthetic syllables in which the vowel could have two standard trajectory shapes. The stimuli were described to the subjects as being dVd synthetic syllables, and the subjects were asked to identify the vowel in the stimuli as any vowel of the vowel system of their language. The subjects were also asked to note whether the perception of the consonant of the dVd syllable disturbed the perception of the vowel.

All the stimuli were synthesized with the Klatt synthesizer (Klatt, 1980). The F1 trajectory was the only parameter by which two stimuli, with the same F1 maximum, differed; depending on the shape of the F1 trajectory, the stimuli were identified as of type I or II (Fig. 2). Two sets of stimuli were used. In the first set (Fig. 2a), the duration of the stimuli was 115

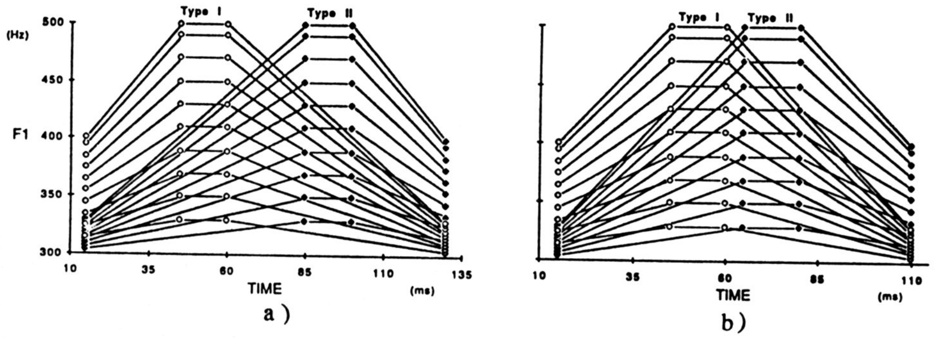

Figure 2. Schematic F1 trajectory for the stimuli in the first set (Fig. 2a) and in the second set (Fig. 2b). The stimuli were of type I and of type II, depending on the two different trajectory shapes, as shown on the figure.

msec and of the steady-state 15 msec, for all stimuli, while the onglide duration was 30 msec for type 1 and 70 msec for type II stimuli. In the second set (Fig. 2b), the duration of the stimuli was 95 msec (20 msec shorter than the stimuli of the first set) and of the steady-state was 15 msec, for all stimuli, while the onglide duration was 30 msec for type I and 50 msec for type II stimuli. In each set, 10 stimuli of each type were considered, with an F1 maximum ranging from 330 to 500 Hertz in steps of 20 Hertz. Four subjects participated in the experiment; two were native speakers of American-English, and two were native speakers of Italian. All the subjects had profound knowledge only in their native language. None of the subjects knew about the purpose of the experiment.

Corresponding to each set of stimuli, the experiment consisted of three phases. In the first phase, each type I stimulus was presented ten times. The ten type I stimuli were ordered to make each stimulus follow another only once. In the second phase, the same procedure was used with type II stimuli. In the third phase, stimuli of both types were presented. On the whole, each stimulus was presented twenty times. At the end of each phase, the subject could rest for few minutes. The test was approximately 45 minutes long.

3. Results of the experiment

The vowels of the stimuli were identified as [i,I,e,ε] by the American subjects and as [i,e,ε] by the Italian subjects. The American subjects characterized the vowel [e] as the non-diphthongized [e^y]. None of the subjects declared that the perception of the consonant of the stimuli (perceived as [d]) disturbed the perception of the vowel.

3.1. *American subjects*

Figure 3 shows the results obtained for each of the two subjects separately, with the first set (long stimuli). Figure 4 shows the results obtained for each subject separately, with the second set (short stimuli). These figures show that stimuli with the F1 maximum in the low frequency range were mainly associated with the tense vowel [i]. Very possibly this was due first to the low value of the F1 maximum and secondly to the very slight difference between the F1 maximum and the onset and offset values (flat shape of the F1 trajectory). Stimuli with the F1 maximum in the medium frequency range were perceived as [I] if of type II, while there was an uncertainty area for type I stimuli (perceived as [i], [I] or [e]). This uncertainty was more evident with short than with long stimuli. In the high fre-

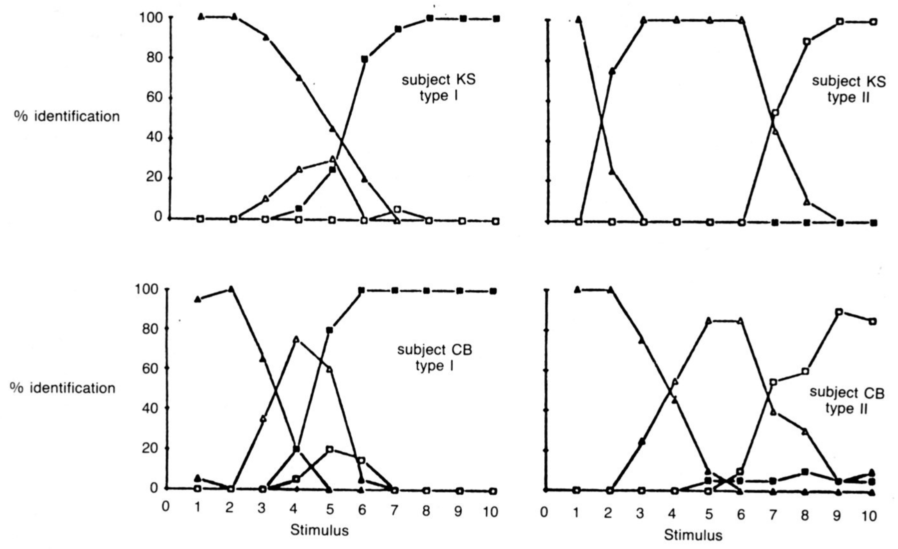

Figure 3. Results of the experiment with the stimuli of the first set (long stimuli) in terms of the percent identification of the tense vowels [i] and [e] vs the lax vowels [I] and [ɛ], for each American subject. ▲: [i], △: [I], ■ [e], □ [ɛ].

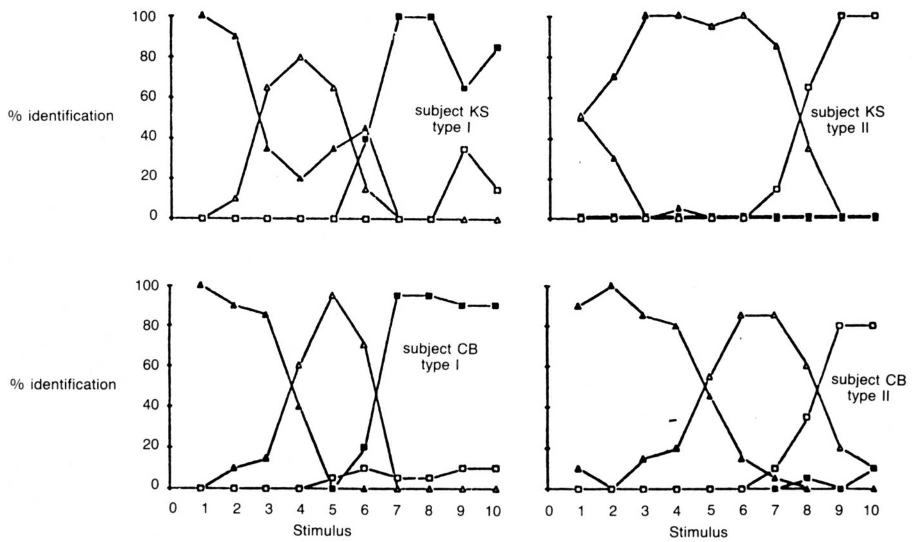

Figure 4. Results of the experiment with the stimuli of the second set (short stimuli) in terms of the percent identification of the tense vowels [i] and [e] vs the lax vowels [I] and [ɛ], for each American subject. ▲: [i], △: [I], ■ [e], □ [ɛ].

quency range, the distinction between the perception of type I and type II stimuli was sharp: type I stimuli were perceived as [e] and type II stimuli as [ɛ].

3.2. *Italian subjects*

The results obtained with the first set (long stimuli) are presented in Fig. 5, which shows that type I stimuli were mainly perceived as the tense vowels [i,e] and type II stimuli as [i] when F1 maximum was in the low frequency range and as [ɛ] when it was in the high range.

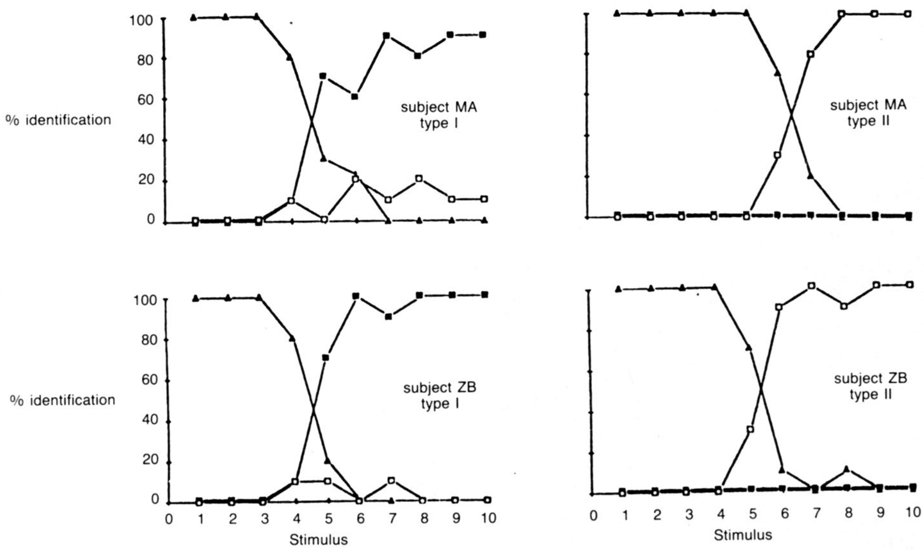

Figure 5. Results of the experiment with the stimuli of the first set (long stimuli) in terms of the percent identification of the tense vowels [i] and [e] vs the lax vowel [ɛ], for each Italian subject. ▲: [i], ■ [e], □ [ɛ].

The results of the experiment with the second set (short stimuli) are shown on Fig. 6. As before, type I stimuli were in general perceived as [i] and [e]. Type II stimuli were in general perceived as [i] when F1 maximum was in the lower part of the frequency range and as [ɛ] when it was in the higher range. Figure 6 shows that the same effect was observed for type II stimuli of the two sets, except that shorter type II stimuli with high F1 maximum were sometimes perceived as [e], while this did not occur with the longer stimuli. Possibly this effect was due to the fact that in the shorter stimuli of type II the F1 maximum was reached more rapidly than in the

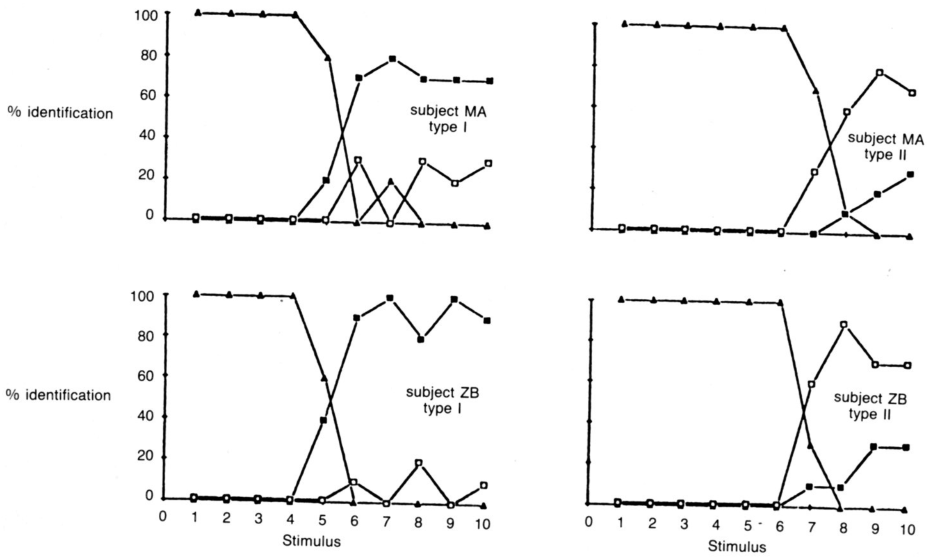

Figure 6. Results of the experiment with the stimuli of the second set (short stimuli) in terms of the percent identification of the tense vowels [i] and [e] vs the lax vowel [ɛ], for each Italian subject. ▲: [i], ■ [e], □ [ɛ].

longer stimuli of type II. Thus, the shorter type II stimuli had the F1 trajectory shape more similar to the F1 trajectory shape of type I stimuli than in longer stimuli. Consequently, the subjects could have shown a tendency to perceive short type II stimuli as tense vowels.

Note that the Italian vowel system does not include the vowel [I], so that the vowel areas of [i] and [e] are contiguous in the F1 dimension; this may have caused the Italian subjects responses not to show the same uncertainty found for the American subjects, when F1 maximum was in the medium frequency range.

4. Discussion

The results of the experiments presented showed that in general stimuli characterized by a fast rising of the F1 trajectory tended to be associated with tense vowels while stimuli with a slow F1 rising were in general perceived as lax vowels, except when the F1 maximum values were low and consequently the F1 trajectory shape was almost flat. For type I stimuli, there was an uncertainty area in the responses of the American subjects, when F1 maximum was in the medium frequency range.

Huang (1985) investigated the influence of the relative duration of onglide, steady-state and offglide on the perception of the tense/lax distinction of American English vowels. The synthetic stimuli considered in the experiment carried out by Huang had three different F1 and F2 trajectories; standard (onglide time: 20 msec, steady-state: 195 msec, offglide time: 20 msec), tense (onglide time: 50 msec, steady-state: 135 msec, offglide time: 50 msec) and lax (onglide time: 50 msec, steady state: 85 msec, offglide time: 100 msec). Longer offglide duration was chosen for lax stimuli, according to Lehiste and Peterson (1961). In fact, Lehiste and Peterson noted, on the basis of acoustic measurements made on monosyllabic CVC words, that for tense vowels the duration of the steady-state was longer than for lax vowels and that for tense vowels the offglide duration was shorter than for lax vowels. Huang's findings were that vowel responses shifted towards tense vowel judgements when stimuli with tense formant trajectories were considered, in the case of either tense/tense or tense/lax non-high vowel pairs. This shift was present also going from the tense to the standard stimuli for non-high vowel pairs. On the results of Huang's experiment, we observed that the shift obtained going from the lax to the tense trajectory stimuli was similar to that found going from the tense to the standard trajectory stimuli.

In another experiment, Huang investigated the role of duration in the perception of vowels of American English. Huang considered stimuli with similar onglide and offglide durations (20 msec) but different steady-state durations (0,50,100,195 msec). Huang's findings were that a shift towards tense vowel judgements were found going from the shortest to the longest duration. A shift was present also in the case of lax/lax pairs, and it was in the direction of the vowel characterized by a higher value of F1.

The results of the experiments of the present study showed that stimuli characterized by a long offglide and a short onglide of F1 tended to be perceived as tense vowels while stimuli with a short offglide and a long onglide tended to be perceived as lax vowels and that shortening the stimuli had little influence on the perception of the synthetic vowels by the American subjects. The divergencies found with the studies mentioned above could be explained as follows: 1) different components were considered in the present study (same steady-state duration and different onglide and offglide durations), and 2) the difference in duration between the stimuli in the two sets was 20 msec and the results of Huang's experiment showed that in the case of the [*i*]-[I] pair the shift was evident going from the shortest to the longest stimuli but that it was small going from the 90 msec to the 140 msec stimuli (which were similar in duration to the stimuli considered in the experiments of the present study). The shift of judgements found by Huang in the case of tense/standard stimuli could be interpreted, on the basis of the results of the present study, as due to the different initial slope of the F1 trajectory. However, this explanation is not unique; an alternative interpretation could be that the determining factor of the effect is the duration of the steady-state, which assumes different values in the two cases. Both interpretations could be valid. In addition in Huang's study the F2 values of the synthetic vowels considered varied along a continuum between adjacent vowels, while in the present study the F2 values were the same for all stimuli.

Further work should tend to specify acoustic correlates of the tense/lax distinction which would be possibly related to the effect observed in the perceptual experiments described. If this link existed, it would be possible to hypothesize that these properties may be used by the auditory system to encode the information about tenseness or laxness of vowels.

ACKNOWLEDGEMENTS

My sincere thanks go to Giovanni Flammia for the numerous helpful suggestions and discussions during the preparation of this manuscript.

REFERENCES

Brozovic D., 1967, "Sull'inventario dei fonemi serbocroati e i loro tratti distintivi". *WSI*, 12, 161-172.

Di Benedetto M.G., 1987, *An acoustical and perceptual study on vowel height*, PhD dissertation, Università degli studi di Roma.

Huang C.B., 1985, *Perceptual correlates of the tense/lax distinction in general American English*, Master's thesis, Massachusetts Institute of Technology, Cambridge (MA).

Lehiste I., Peterson G.E., 1961, "Transitions, glides and diphtongs". *J. Acoust. Soc. Am.* 33, 3, 268-277.

Klatt D.H., 1980, "Software for cascade/parallel formant synthesizer". *J. Acoust. Soc. Am.* 67, 3, 971-995.

Muljacic Z., 1972, *Fonologia della lingua italiana*, Bologna, p. 29.

Saltarelli M., 1970, *A phonology of Italian in a generative grammar*, L'Aia, pp. 40-92.

Stevens K.N., Keyser S.J., Kawasaki H., 1986, "Toward a phonetic and phonological theory of redundant features". In: Perkell J., Klatt D. (eds.), *Symposium on Invariance and Variability of Speech Processes*, Lawrence Erlbaum Assoc., Hillsdale, NJ, pp. 426-449.

Tonelli L., 1984, *Uberlegungen zur naturalichen phonologie: kontrastive analyse des Bosnerischen und des Sudtiroler Italienische*, Serie: Monografie-N. 23, Trieste.