TRADING RELATIONS BETWEEN CUES FOR THE PHARYNGEALIZED/ NON PHARYNGEALIZED CONTRAST

Mohamed YEOU
Institut de phonétique, Sorbonne Nouvelle. CNRS , Paris.

ABSTRACT

The perceptual effects of orthogonal variations in two acoustic parameters (F1 and F2 onset frequencies) which differentiate Arabic pharyngealized /sˤ/ from plain /s/ were examined. An identification task showed a systematic displacement of the perceptual boundary as the onset value of F1 ($F1_0$) changes from low (250 Hz) to high (460 Hz), thus reflecting a trading relation between the two cues ($F1_0$ and $F2_0$). To investigate whether or not discrimination accuracy was differentially affected by the phonetic cooperation or conflict between the two cues, an AX discrimination task was used. As predicted, the discriminability ordering was the following: cooperating cues > one-cue > conflicting cues.

INTRODUCTION

In Arabic the four consonants /ð, t, d, s/ have the corresponding following pharyngealized consonants /ðˤ, tˤ, dˤ, sˤ/. These latter have in addition to a primary articulation (dental/alveolar contact), which they share with the former, a secondary articulation (backing of the tongue towards the pharyngeal wall). The acoustic consequence of this double articulation is a considerable lowering of F2 and a slight raising of F1 in vowels adjacent to pharyngealized consonants. In [1] locus equations which encode the dynamics of the F2 transition were capable of distinguishing pharyngealized consonants from non-pharyngealized ones. The purpose of this study is to investigate which acoustic properties identify pharyngealized consonants.

2. INDENTIFICATION TASK

2.1. Method

2.1.1. Stimuli

Four series of [s-sˤ] continua, each in the vowel context of [i:] were generated using a software parallel synthesizer [2]. Formant frequency values and timing characteristics for the [si:]-[sˤi:] series were adapted from the average values for a male Moroccan native speaker of Arabic. Two reference stimuli were used in the experiments: a pharyngealized type and a non-pharyngealized type. For the first type, F1 and F2 onset frequencies were 460 Hz and 1060 Hz, respectively. For the second type, the onset frequencies were 250 Hz and 1800 Hz. As Figure 1 shows the continua were constructed by systematically varying these onset frequecies in 10 steps of 35 Hz for F1 and in 10 steps of 140 Hz for F2.

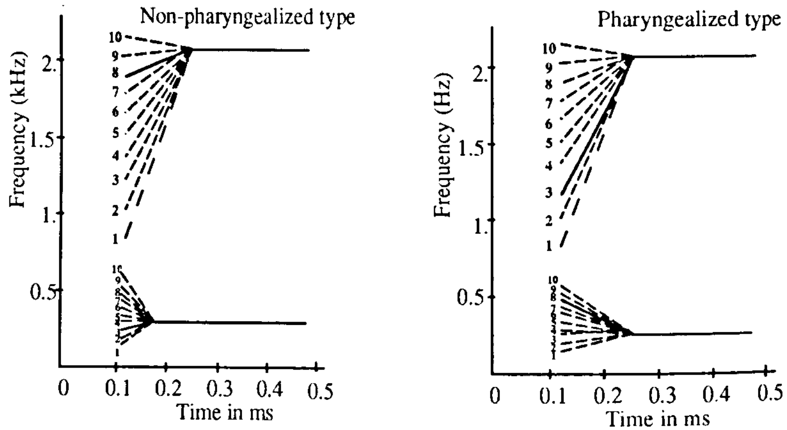

Figure 1. Schematic trajectories showing variations in $F1_0$ and $F2_0$ for the non-pharyngealized type continua (left) and pharyngealized-type continua (right). Reference stimuli are presented by solid lines.

All stimuli were of 470-ms duration and contained the same FO. Figure 1 provides a schematic representation of the continua: phayngealized and non-pharyngealized types.

2.1.2. Procedure and Subjects

The identification test consists of a randomization of the stimuli from the four series of continua. Each stimulus was presented 5 times. Intervals between stimuli were 2.5 s and between ten-stimulus blocks were 9 s.
The subjects were 11 graduate students (phonetics/linguistics). All subjects are Moroccan native speakers of Arabic and reported having normal hearing.

2.2. Results

Group identification functions for the F1 continua (both the pharyngealized type and the non-pharyngealized one) are presented in Figure 2. These functions show that variations in the onset F1 frequency are not effective in producing a perceived contrast between /sˤi:/ and /si:/: no crossing of boundaries occurs.

Figure 3 displays group identification functions for the F2 continuum (both pharyngealized and non-pharyngealized types). The functions show that the onset frequency of F2 is a critical acoustic property for the perception of the pharyngealized/non-pharyngealized contrast. Category boundaries were evaluated by interpolating the stimulus number at the 50 % crossover. The boundary is at 1276 Hz (near stimulus 4) for the pharyngealized-type continuum, and at 1773 Hz (near stimulus 8) for the non-pharyngealized-type continuum. The important difference between the two boundaries indicate that subjects would need an additional F2 onset lowering of 497 Hz to begin hearing [sˤi:] when F1 onset frequency is not appropriate for [sˤi:]. A t-test shows that such boundary difference is significant, $t(109)= 11.6, p < 0.0001$.

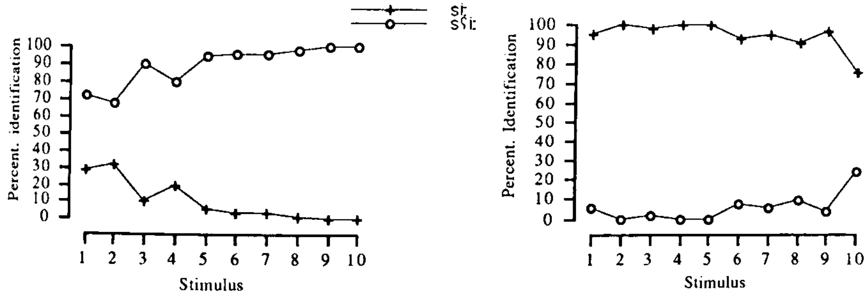

Figure 2. Identification functions for the F1 continua: pharyngealized type (left) and non-pharyngealized type (right).

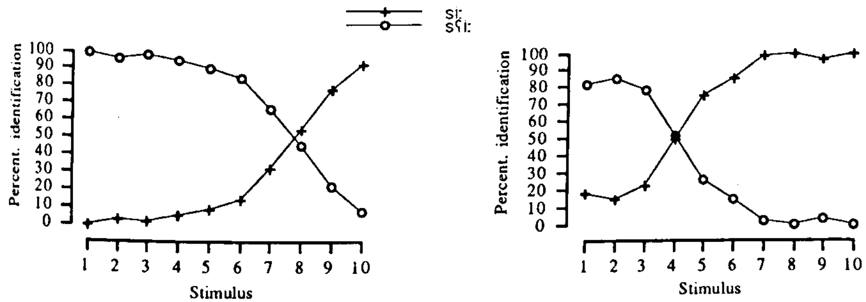

Figure3. Identification functions for the F2 continua: pharyngealized type (left) and non-pharyngealized type (right).

3. DISCRIMINATION TASK

3.1. Stimuli and procedure

Evidence for a trading relation between $F1_0$ and $F2_0$ was derived from the identification results of experiment 1. This would imply a perceptual equivalence between the two cues. The second experiment used a method testing the possibility of such perceptual equivalence [3]. This method investigates whether discimination performance would be differentially affected by the cooperation or conflict of the two cues along the phonetic dimension. Three comparing conditions of cue combination were used for this purpose: (1) one-cue condition, in which only $F2_0$ was varied; (2) cooperating two-cue condition, in which both $F1_0$ and $F2_0$ complemented each other phonetically (one member of each pair had one cue biased towards [si:], the other toward [sˤi:]); and (3) conflicting two-cue condition, in which the two cues cancelled each other. All ten stimuli of the continua were utilized, and each stimulus was paired with stimuli which were 4 steps apart from it on the spectral dimension (Δ 560 Hz). This amount approximates the amount of the boundary shift found in the identification experiment (497 Hz). The discrimination test (an AX task) was a randomised sequence of all possible stimulus comparisons repeated 5 times. The interval within each pair was 0.5 s and between succesive pairs 2.5 s.

3.2. Results

Group predicted discrimination scores derived from the identification data[1] are presented in Fig 4 and the corresponding obtained discrimination scores in Fig 5. A repeated-measure ANOVA with Conditions of comparison X Scores (predicted vs. obtained) was conducted on the results. There was no performance difference betwen predicted and observed scores across the cooperating cues condition, $F(1,131)=0.380$, $p=0.53$, and across the conflicting cues condition, $F(1,131)=0.345$, $p=0.55$. There was, however, a small significant effect for observed vs predicted scores across the one cue condition, $F(1,131)=5.390$, $p<0.02$. This difference indicates some ability to discriminate acoustic differences on a non-phonetic basis. This is also revealed by the fact that under the cooperating cues condition, the boundary-related peak was not well marked in the obtained scores compared with the predicted scores.

The above discrepancies are not damaging to the perceptual equivalence hypothesis, which is basically confirmed if discrimination performance in the cooperating cues condition is higher than performance in the conflicting cues condition. An ANOVA crossing Types of comparisons with Stimuls pairs was performed on the obtained data. Overall differences between types of comparisons were significant, $F(2, 180)= 188.93$, $p<0.0001$, and *post hoc* comparisons (*Fisher PLSD*) supported the perceptual equivalence prediction that the discriminability ordering would be: cooperating cues > one cue > conflicting cues.

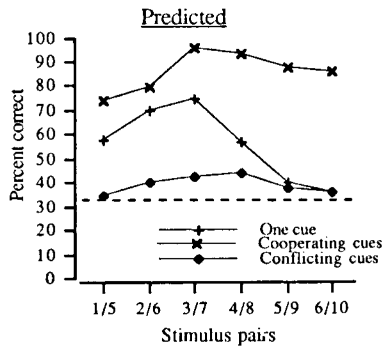

Figure 4. Group predicted discrimination scores in function of stimulus pairs and conditions of cue combination.

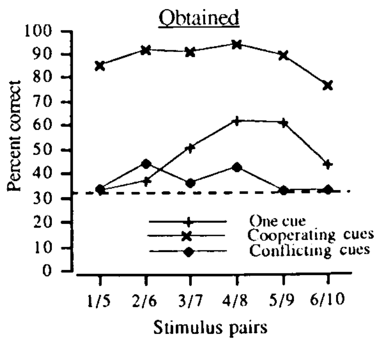

Figure 5. Group obtained discrimination scores in function of stimulus pairs and conditions of cue combination.

DISCUSSION

Most previous studies on trading relations have involved cues which are acoustically dissimilar and temporally separate and have generally supported the hypothesis that trading relations reflect phonetic perception which 'refers' to articulation/production [3, 4, 5, 6].

The present study replicates the findings of these studies for a new contrast involving two cues both spectral and temporally co-occuring. This would suggest the possibility of an interaction of some psychoacoustic origin, similar to that reported by [6,7], where the cues investigated consituted the same portion of the signal. The question that now arises: what is the phonetic and the auditory origins of the intergration of $F1_0$ and $F2_0$.

The phonetic explanation is based upon the hypothesis which explains trading relations with reference to articulation. Accordingly, the two cues are perceptually integrated because they are the result of the same articulatory gesture, i.e. pharyngealization. Such an articulatory rationale is not impossible; it is the most straighforward. In production of a pharyngealized consonant, the coarticulated /i:/ shows a high $F1_0$ (460 Hz) and a low $F2_0$ (1040 Hz). Both acoustic events are the result of a unitary articulation maneuver which includes: (1) a rearward movement of the back of the tongue towards the pharyngeal wall; and (2) a depression of the tongue's palatine dorsum [8]. The two movements result in a widened oral cavity and a reduced pharyngeal cavity. The high onset of F1 seems to be due to the reduction of the pharyngeal cavity, while the low onset of F2 to the widening of the oral cavity.

The psychoacoustic explanation is based on an auditory coherence account which proposes that listeners perceive the speech patterns of speech according to Gestalt principles. The principle that concerns us here is that of temporal proximity. $F1_0$ and $F2_0$ may cohere by virtue of their temporal proximity. They both bave onsets and offsets that are temporally simultaneous, i.e. they have the same duration (80-120 ms in the case of the vowel /i:/). Moreover they are very close in their frequencies with only a 580-Hz distance apart. This explanation is similar to that given in [9], where the harmonics of a vowel formant cohere by virtue of their temporal proximity.

[1] The following formula was used: $P_{corr}=[1+2(P_a-P_b)^2]/3$, where P_{corr} is the predicted probability of correct responses for a given stimulus pair, P_a is the obtained [sˤi:] responses to stimulus *a*, and P_b is the obtained [sˤi:] responses to stimulus *b* in the comparison. Chance level is at 0.33.

REFERENCES

[1] Yeou, M. (1995) "An investigation of locus equations as a source of information for consonantal place of articulation," *4th European Conference on Speech Comm. and Technology,* Madrid.
[2] Klatt, D.H. (1980) "A software for a cascade/parallel formant synthesizer," *JASA* 67: 971-995.
[3] Fitch, L., T. Halwes, & D.M. Erickson. (1980) "Perceptual equivalence of two acoustic cues for stop-consonant manner," *Perception and Psychophysics* 27 (4): 343-350.
[4] Best, C.T., B. Morrongiello, & R. Robinson (1981) "Perceptual equivalence of acoustic cues in speech and nonspeech perception," *Perception and Psychophysics* 29: 191-211.
[5] Repp, B.H. (1982) "Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception," *Psychological Bulletin* 92: 81-110.
[6] Repp, B.H. (1983) "Trading relations among acoustic cues in speech perception are largely a result of phonetic categorization, " *Speech Communication* 2: 341-362.
[7] Polka, L. & Strange, W. (1985) "Perceptual equivalence of acoustic cues that differentiate /r/ and /l/," *JASA* 78: 1187-1197.
[8] Ghazeli, S. (1977) *Back Consonants and Backing Coarticulation in Arabic.* Ph.D. Dissertion, Texas University at Austin.
[9] Darwin, C.J. (1984) "Perceiving vowels in the presence of another sound: Constraints on formant perception," *JASA* 76: 1636-1647.