355

LANGUAGE AND SPEECH, Vol. 24, Part 3, 1981

ACOUSTIC FACTORS THAT MAY CONTRIBUTE TO CATEGORICAL PERCEPTION*

JANET G. MAY

Haskins Laboratories, New Haven, Connecticut

The perception of the voiced and voiceless velar and pharyngeal fricatives /ɣ, x, ʕ, ħ/ and of /ʃ, s/ in Colloquial Egyptian Arabic was examined to determine if the presence of the first two or three formants in /ɣ, x, ʕ, ħ/ results in continuous perception, in contrast to an expected categorical perception of /ʃ, s/, which lack these formants. Three twelve-step series of VFV nonsense words were synthesized. For the /ʃ/–/s/ series, the center of a band of high-frequency noise was varied in equal steps. For the /x/–/ħ/ and /ɣ/–/ʕ/ series, the first formant was varied. Eight native speakers were asked to identify the stimuli and discriminate two-step differences in a 4IAX discrimination task. While the voiced /ɣ/–/ʕ/ series showed continuous or less categorical perception than the /ʃ/–/s/ series, the voiceless /x/–/ħ/ series was perceived somewhat categorically. This suggests that voicing alone, or in combination with acoustic information about the lower formants, may be a necessary condition for continuous perception.

INTRODUCTION

Although the past thirty years have witnessed a revolution in speech research, one of the earliest discoveries made about speech perception still remains somewhat of a mystery: the finding that some speech sounds are perceived in a manner quite different from others. Stop consonants are usually perceived categorically: Subjects can only discriminate as many sounds as they have different labels for (Liberman, Harris, Hoffman, and Griffith, 1957). On the other hand, vowels are perceived more or less continuously: Subjects can discriminate acoustic differences between phonetically equivalent stimuli (Fry, Abramson, Eimas, and Liberman, 1962).

Categorical perception, however, is not speech-specific (see Strange and Jenkins, 1978). It has been demonstrated for such psychophysical continua as noise-buzz sequences, tone onset times, and visual flicker fusion (Miller, Weir, Pastore, Kelly, and Dooling, 1976; Pisoni, 1977; Pastore, Ahroon, Baffuto, Friedman, Puleo, and Fink, 1977). In addition, the degree of categorical perception can be manipulated by training, experience, task variables, interstimulus relations, and other experimental factors. For example,

---

* *This paper is based upon a 1979 University of Connecticut doctoral dissertation entitled "The Perception of Egyptian Arabic Fricatives." A shorter version of this paper was presented at the 97th Meeting of the Acoustical Society of America, Boston, Spring 1979. This research was supported by PHS Grant 00321-05 to the University of Connecticut and the preparation of this manuscript was supported by NICHD Grant HD-01994 and BSR Grant RR-05596 to the Haskins Laboratories. Many thanks to Ignatius G. Mattingly, Arthur S. Abramson, and Bruno H. Repp for their helpful advice at all stages of work.*

subjects can be trained to perceive voicing and place features in stop consonants non-categorically (Barclay, 1972; Carney, Widin, and Viemeister, 1977; Samuel, 1977). If vowels are shortened, put in CVC syllables, or degraded by adding noise to them, they show a tendency for categorical perception (Lane, 1965; Stevens, 1968; Sachs, 1969; Fujisaki and Kawashima, 1968); furthermore, increasing the interstimulus interval will cause an increase in the degree of categorical perception (Pisoni, 1971).

To account for the perceptual difference between stop consonants and vowels, Fujisaki and Kawashima (1969; 1970; Fujisaki, 1979) proposed a model of speech perception in an experimental situation. They suggested that when a subject hears a speech stimulus, he or she stores two kinds of information about it in short term memory: an echoic memory containing information about the acoustic details of the sound, and a phonetic memory containing a phonetic label. Due to its discrete nature, phonetic memory will endure longer than echoic memory. Furthermore, since stop consonants are short, their echoic memories will decay rapidly, and therefore may not be available to enable a subject to discriminate phonetically equivalent stimuli. Consequently, he or she will have to refer to labels stored in phonetic memory that will allow discrimination of only as many stimuli as the subject has different labels for. Since vowels are much longer in duration, their echoic memories will persist longer than those for stops, and will probably be available when a subject needs them. The information in echoic memory will allow the subject to discriminate acoustic differences between phonetically equivalent stimuli. This would explain why stops are perceived categorically and why vowels are perceived continuously.[1]

If, indeed, long duration is a necessary condition for continuous perception, it is certainly not a sufficient conditon. The fricatives /ʃ/ and /s/, which can have durations comparable to those of vowels, are perceived categorically (Fujisaki and Kawashima, 1968; 1969; Repp, in press). In the production of /ʃ/ and /s/, free zeros created by the

---

[1] *There is corroborative evidence for the existence of echoic memory from tests of immediate ordered recall of auditorily presented consonants and vowels. It is assumed that a subject must hold acoustic information about the stimuli in a sensory or pre-linguistic form for at least a few seconds until it can be analyzed. This store was termed Precategorical Acoustic Storage (PAS) by Crowder and Morton (1969), and is equivalent to echoic memory. Crowder (1971) found that when subjects are asked to recall a series of vowels, they show a significant improvement on the last few members of the series. This recency effect was attributed to the existence of PAS for most recently received vowels, which acts to improve their recall. Since PAS lasts only a few seconds, the PAS for the earlier members of the series will have decayed by the time the subjects are required to recall the series. In additon, when a verbal suffix, which subjects are told to ignore, is added to the end of the series, it seems to interfere with the PAS of vowels and the recency effect is lost. This suffix effect was attributed to interference of the suffix with the PAS of the most recent vowels. It is very interesting to note that neither the recency effect nor the suffix effect was found for the voiced stop consonants. Since stops are relatively short in duration, their PAS may not endure as long as that for vowels. Therefore, the PAS of stop consonants will not be available and so cannot help to improve recall of the last items in the consonant series. In addition, whatever information is available in PAS will be overwritten by a suffix.*

cavity behind the constrictional source cancel the lower formant frequencies from the spectra of these fricatives. Perhaps the absence of these formants causes categorical perception by somehow making the echoic memory unreliable, and therefore not available to the subject.

Fricative consonants of Egyptian Arabic

Colloquial Egyptian Arabic offers the opportunity to test this hypothesis, since its phonetic inventory contains fricatives produced in both the front and back cavities of the vocal tract. The front cavity fricatives are the familiar /ʃ/ and /s/. The back cavity fricatives are the less familiar voiced and voiceless velars /ɣ/ and /x/, as in /ɣɛnɛm/ 'sheep' and /xaʃi/ 'animal's mouth', respectively, and the voiced and voiceless pharyngeals /ʕ/ and /ħ/, as in /ʕans/ 'goat' and /ħomar/ 'donkey', respectively. In the production of these back cavity fricatives, the constrictional source is close to the glottis, making the cavity behind the source very short. Such a tube produces anti-resonances with frequencies too high to zero out the lower formants. It was hypothesized that the presence of distinctive lower formants would allow continuous perception of these fricatives by making their echoic memories more dependable.

Scholars generally agree that /x/ and /ɣ/ are velar fricatives, although some describe them as uvular, and that /ħ/ is a pharyngeal fricative (Gairdner, 1925; Harrell, 1957; Khalafallah, 1969; Aboul-Fetouh, 1969; Abdel-Malek, 1972). But, while the history of /ʕ/ attests to its phonological status as the voiced counterpart of /ħ/, there may be some question about its phonetic status as a voiced fricative. Laufer and Condax (1981), using a fiberscope, have found that both of these consonants are produced in a dialect of another Semitic language, Hebrew, with a closure or constriction of the epiglottis at the pharyngeal wall; this technique should be used on Egyptian Arabic as well. In the present study, in intervocalic position /ʕ/ appeared as a glide-like continuation of the adjacent vowels' formants, occasionally with some frication, and in initial position sometimes as a voiceless fricative or a voiceless stop.

Recordings were made of a native speaker of Colloquial Egyptian Arabic producing the fricatives /ʃ, s, x, ɣ, ħ, ʕ/ in intervocalic position. Spectrographic analysis revealed that Arabic /ʃ/ and /s/ are similar to the corresponding English fricatives, /ʃ/ having spectral energy in the 2000–6000 Hz range, and /s/ in the 4000–7000 Hz range (cf. Strevens, 1960). Average formant frequencies for the velar and pharyngeal fricatives in an intervocalic /ə/ environment are given in Table 1. Notice that F1 is approximately 900 Hz for the pharyngeal /ħ/ and /ʕ/, but considerably lower for the velars (F1 of /x/ is even lower in the environment of high vowels, probably due to the influence of their low F1s.). The voiceless fricatives /x/ and /ħ/ have an F2 of about 1900 Hz, while for their voiced counterparts it is about 300-500 Hz lower. F3 is about 2400 Hz for the voiceless and 2200 Hz for the voiced fricatives, however the amplitude of F3 for /x/ is usually very low. Indeed, it was found that this formant was not needed to synthesize acceptable /x/ and /ħ/. Similarly, F4 of /ħ/ was also very low in amplitude and was never visible for /ʕ/; therefore it is not surprising that it was not needed to synthesize acceptable /ħ/ and /ʕ/. For the velar /x/ and /ɣ/ F4 was approximately 3900 Hz.

TABLE 1

Average Formant Frequencies (in Hz) of Arabic Fricatives in an Intervocalic /ə/ Environment

	F1	F2	F3	F4
/x/	700	1850	2400	4050
/ɣ/	450	1550	2250	3700
/ħ/	900	1900	2400	4000
/ʕ/	875	1380	2200	——

METHOD

Stimuli

Three twelve-step series of VFV stimuli were created on a Glace-Holmes terminal analog synthesizer (Glace, 1968). The first was a series from /ʃ/ to /s/, the second from /x/ to /ħ/, and the third from /ɣ/ to /ʕ/.

All stimuli in each series contained the same initial and final /ə/, which was 140 msec long and contained appropriate formant frequency transitions to steady-state segments representing the intervocalic fricatives. In its initial steady-state this vowel had an F1 of 658 Hz, and F2 of 1521 Hz, and an F3 of 2329 Hz.

Each fricative segment in the /ʃ/–/s/ series (Figure 1) was 220 msec long and consisted of a band of high-frequency noise, whose center frequency increased from 2974 Hz for /ʃ/ to 4784 Hz for /s/ in steps of about 165 Hz. Sixty-msec transitions for F1, F2, and F3 occurred in the vocalic segments starting with the vowel's steady-state values and ending with 440, 1845, and 2652 Hz, respectively, for /ʃ/, and 440, 1764, and 2652 Hz, respectively, for /s/. Thus, only the F2 transition varied across the series.

Each fricative segment in the /x/–/ħ/ series (Figure 2) was 200 msec long and consisted of the first two noise-excited vocalic formants and a band of high-frequency noise. For all stimuli the second formant was 1886 Hz, and the center of the band of noise was 3961 Hz. The first formant increased from 368 Hz for /x/ to 900 Hz for /ħ/ in steps of about 50 Hz. The amplitude of the high-frequency noise decreased from –24 dB (with respect to the amplitude of the vowel's first formant) for /x/ to –39 dB for /ħ/. Thirty-msec transitions for F1, F2, and F3 occurred in the vocalic segments starting with the vowel's steady-state values and ending with 465, 1764, and 2248 Hz,

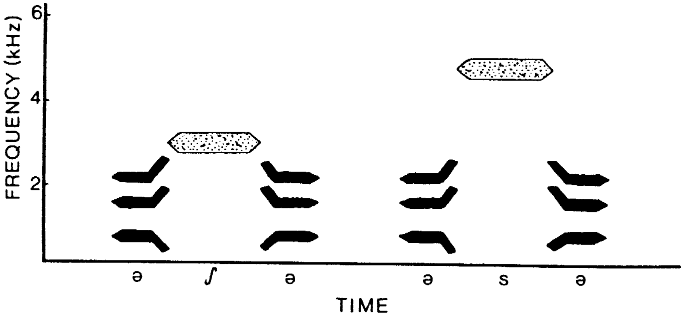

Fig. 1. Schematic spectrograms of stimulus #1 (on left) and of stimulus #12 (on right) in the synthetic /ʃ/—/s/ series.

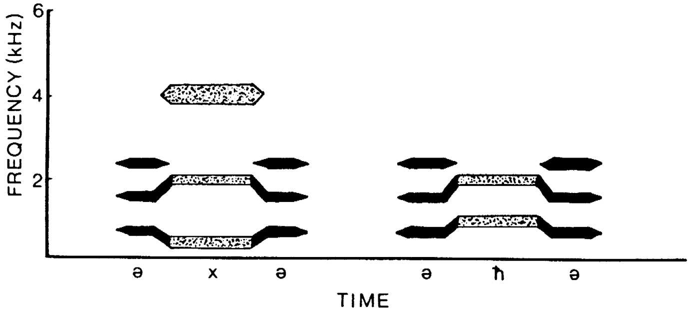

Fig. 2. Schematic spectrograms of stimulus #1 (on left) and of stimulus #12 (on right) in the synthetic /x/—/ħ/ series.

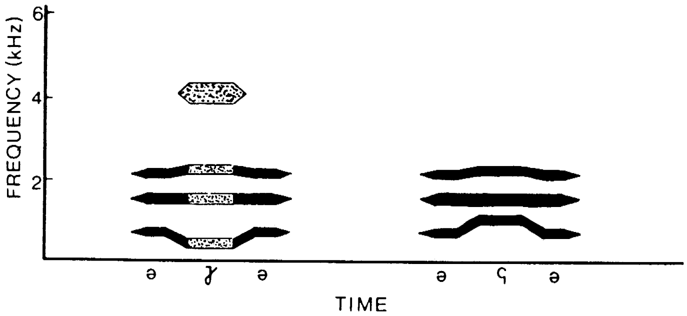

Fig. 3. Schematic spectrograms of stimulus #1 (on left) and of stimulus #12 (on right) in the synthetic /ɣ/—/ʕ/ series.

respectively, for /x/, and 827, 1764, and 2248 Hz, respectively, for /ħ/. Thus, only the F1 transition varied across this series.

Each fricative segment in the /ɣ/–/ʕ/ series (Figure 3) was 110 msec long, and consisted of three vocalic formants and a band of high-frequency noise. For all these segments the second formant was 1521 Hz, the third formant, 2248 Hz, and the center of the band of noise, 3961 Hz. The first formant increased from 368 Hz for /ɣ/ to 900 Hz for /ʕ/ in steps of about 50 Hz. The amplitude of the high-frequency noise was decreased from –13 dB for /ɣ/ to –39 dB for /ʕ/. The vocalic formants and the band of noise were synthesized using a mixture of periodic and aperiodic excitation. The ratio of periodic to aperiodic excitation increased with each step along the series. This was achieved by interspersing an increasing number of 10-msec intervals of periodic excitation among a decreasing number of 10-msec intervals of aperiodic excitation, until the last stimulus in the series contained only periodic excitation during this segment. Fifty-msec transitions for F1, F2, and F3 occurred in the vocalic segments starting with the vowel's steady-state values and ending with 416, 1521, and 2248 Hz, respectively, for /ɣ/, and 852, 1521, and 2248 Hz, respectively, for /ʕ/. Thus, only the F1 transition varied across this series.

Experimental tests

One identification test and three 4IAX discrimination tests were prepared for each series of stimuli. In the identification test, subjects were asked to identify each of the 12 stimuli in a series 16 times, by writing down the Arabic letter corresponding to the sound they heard. In each of the three discrimination tests subjects were asked to discriminate each two-step difference eight times, totaling 24 trials across the three tests. The odd stimulus occurred in each position of the 4IAX pairs an equal number of times. A subject responded by writing "1" or "2" to indicate whether the first or second pair of stimuli contained different sounds.

Subjects

Eight phonetically naive adult native speakers of Egyptian Arabic (not including the original native informant), all from Cairo or nearby, were used as paid subjects in these experiments. One subject showed somewhat erratic behavior in the /ʃ/–/s/ identification test, although her discrimination curves for this series showed a peak where one would expect a phoneme boundary. Since discrimination performance predicted from these identification data would be rather irregular, it would be difficult to compare it to the obtained discrimination. In addition, results from most other tests indicate that she was generally an inattentive subject. Consequently, this subject was eliminated from the study.

Procedure

Each subject took 12 tests: one identification test and three discrimination tests for each of the three continua. The subjects were first given all four tests for the /ɣ/–/ʕ/ series, then all tests for the /ʃ/–/s/ series, and finally all tests for the /x/–/ħ/ series. The subjects were divided into two groups of four. Within each group of four tests for a given series, one group of subjects always heard the identification test first, while the other group heard the discrimination tests first. Two tests were administered per experimental session: either one identification test and one discrimination test, or two discrimination tests. Each test took approximately 15 minutes. The subjects had a brief rest period between the two tests. Their responses for the /ʃ/–/s/ series were very inconsistent. Presumably, this was caused by "clipping" of the signal due to a rather high playback level. Therefore, after all other tests had been administered, the /ʃ/–/s/ identification and discrimination tests were presented to subjects with a reduced playback level for a second time. The results of this second presentation are reported here.

RESULTS

Identification

Individual responses were sufficiently alike to warrant pooling of the data. Pooled identification functions are shown in the top halves of Figures 4–6. Each point represents 112 judgments, 16 per subject. The functions for each of the three series demonstrate that subjects consistently divided each into two discrete categories: /ʃ/ and /s/, /x/ and /ħ/, or /ɣ/ and /ʕ/.

Discrimination

Comparison of the group that took all identification tests first with the group that took all discrimination tests first showed that there was no statistically significant difference (Student's *t* test) between the two groups in the discrimination performance for each of the three continua. Therefore, responses from both groups were pooled. In addition, subjects did not exhibit a bias for responding "1" or "2" on any of the discrimination tests (Student's *t* test).

Ideal categorical perception is characterized by a subject's ability to discriminate only as many sounds as he or she can identify, as predicted by the following formula (see Pollack and Pisoni, 1971 for derivation):

$$P(C) = \frac{(a-a')^2 + (b-b')^2 + 2}{4}$$

where P(C) represents the probability of correctly discriminating stimuli A and B, a = P(a | A) (the probability of labeling stimulus A as phoneme *a*), a′ = P(a | B) (the probability of labeling stimulus B as phoneme *a*), b = P(b | A), and b′ = P(b | B).[2]

---

[2] *It has been suggested that categorical perception is characterized not only by predictability, but also by absoluteness – the ability to remain unaffected by surrounding context. Therefore, a more accurate measure of degree of continuous perception would involve comparing obtained discrimination with discrimination predicted from an identification test that used the same context (Repp, Healy, and Crowder, 1979). This procedure was brought to my attention too late to be used in these experiments, but will be used in the future.*

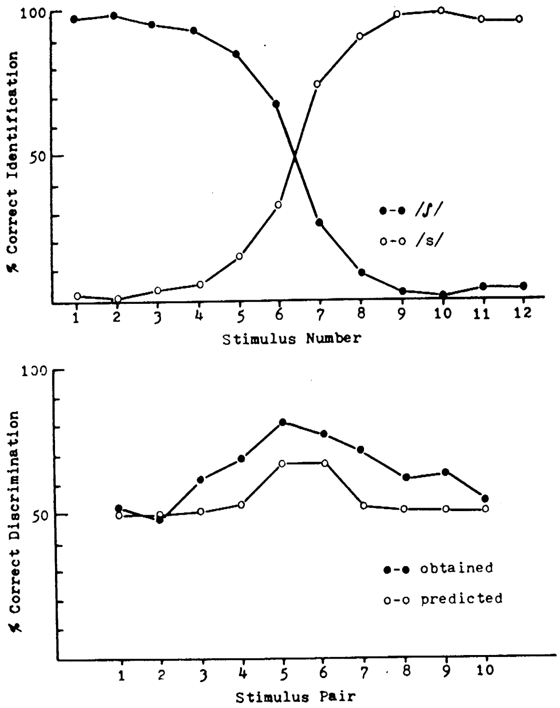

Fig. 4. Identification function (top) and predicted and obtained discrimination functions (bottom) for seven subjects for the /ʃ/–/s/ series.

These predictions are represented in the bottom halves of each of the Figures 4–6 by the open circles. Obtained discrimination scores are denoted by the closed circles, each of which represents 168 judgments on the composite function, 24 per subject. The stimulus pair labeled "1" refers to a pair composed of stimuli 1 and 3, etc.

The identification function in Figure 4 shows that the phoneme boundary for the /ʃ/–/s/ series is located between stimuli 6 and 7. Predicted discrimination shows that, if categorical perception obtains, subjects should not be able to discriminate stimulus pairs 1–4, all of whose members are within the /ʃ/ category, and stimulus pairs 7–10, all of whose members are within the /s/ category (50% = chance). Discrimination performance should increase to about 65% for stimulus pairs 5 and 6 whose members are near the phoneme boundary. Obtained discrimination scores are higher than predicted, $F(1,6) = 16.1$, $p < 0.01$, but show a correlation with predicted discrimination. Note that discrimination performance is greatest for stimulus pairs 5 and 6, as predicted.

The identification function in Figure 5 shows that the phoneme boundary for the /x/–/ħ/ series lies close to stimulus 6. Predicted discrimination shows that, if categorical perception obtains, subjects should not be able to discriminate stimulus pairs 1–3, all of whose members lie within the /x/ category, and stimulus pairs 7–10, all of whose

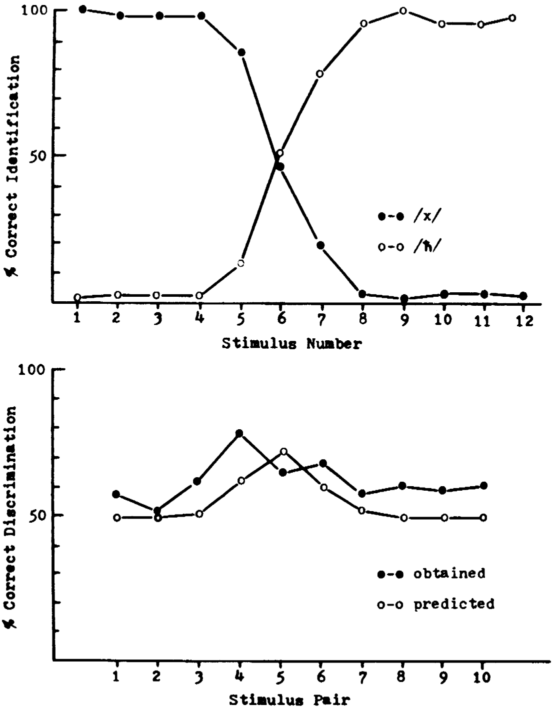

Fig. 5. Identification function (top) and predicted and obtained discrimination functions (bottom) for seven subjects for the /x/–/ħ/ series.

members lie within the /ħ/ category. Discrimination performance should increase to about 72% for stimulus pair 5, whose members, namely 5 and 7, straddle the phoneme boundary. Obtained discrimination, though somewhat higher than predicted, $F(1,6) = 22.6$, $p < 0.005$, shows a correlation with predicted discrimination. Discrimination performance increased from 50–60% for stimulus pairs 1 and 2 to 79% for stimulus pair 4, and then decreased to around 60% for stimulus pairs 7–10. Notice that although performance peaks for stimulus pair 5 in the predicted discrimination, it peaks for stimulus pair 4 in the obtained discrimination. However, the members of both these pairs straddle the phoneme boundary, which is located slightly to the left of stimulus 6.

The identification function in Figure 6 shows that the phoneme boundary for the /ɣ/–/ʕ/ series is located between stimuli 7 and 8. Predicted discrimination shows that for categorical perception, subjects should be able to discriminate stimulus pairs 1–5, all of whose members lie within the /ɣ/ category, and stimulus pairs 8–10, all of whose members lie within the /ʕ/ category, only about 50% of the time. Discrimination performance should increase to about 68% for stimulus pair 6, whose members, namely 6 and 8, straddle the phoneme boundary. Obtained discrimination was significantly greater

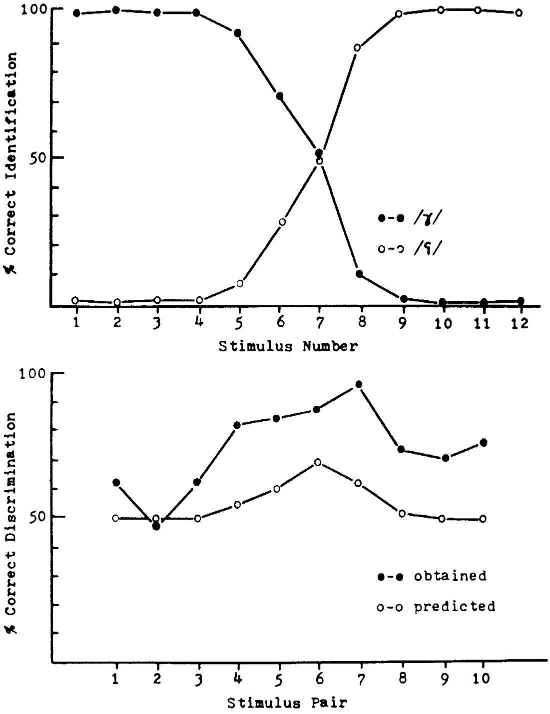

Fig. 6. Identification function (top) and predicted and obtained discrimination functions (bottom) for seven subjects for the /ɣ/–/ʕ/ series.

than predicted discrimination, $F\ (1{,}6) = 142.4, p < 0.001$. Performance increases from about 50% for stimulus pair 2 to about 81% for stimulus pair 4. Performance remains at about 70% for stimulus pairs 4–10, and peaks to about 95% for stimulus pair 7, whose members, namely 7 and 9, straddle the phoneme boundary.

These data demonstrate that subjects tend to perceive the voiceless synthetic stimuli in the /ʃ/–/s/ and /x/–/ħ/ series categorically, while they perceive the voiced synthetic stimuli in the /ɣ/–/ʕ/ series less categorically, or more continuously. An analysis of variance shows this difference to be statistically significant, $F\ (2{,}12) = 12.2, p < 0.005$.

DISCUSSION

The hypothesis examined here is that categorical perception of /ʃ/ and /s/ may, in part, be caused or promoted by a lack of information about the lower formant frequencies in the acoustic signal. It was hypothesized that stimuli in the /ʃ/–/s/ series, which lack these formants, would be perceived categorically, and that stimuli in the /x/–/ħ/ and /ɣ/–/ʕ/ series, which contain these formants, would be perceived continuously. However, the data from the present experiment show that while subjects indeed perceive the voiced fricatives in the /ɣ/–/ʕ/ series continuously, and the voiceless fricatives in the /ʃ/–/s/ series categorically, they tend to perceive the voiceless /x/–/ħ/ series categorically. Since all stimuli are of relatively long duration, it cannot be short duration of acoustic cues that is causing categorical perception in this instance. Although these sounds contain information about the acoustic details of the lower formant frequencies, for some reason the echoic stores seem to be unreliable. As a result, subjects cannot use information stored in them to discriminate stimuli, resulting in categorical perception. It is possible that noncategorical perception requires, in addition to long duration, not only information about the lower formant frequencies, but also that the stimuli be voiced. In fact, the present data could be explained on the basis of voicing alone: The voiceless fricatives /ʃ, s, x, ħ/ were perceived categorically, and the voiced fricatives /ɣ, ʕ/ were perceived continuously (just as vowels).

It is interesting to note that results from experiments involving tests of immediate ordered recall of auditorily presented fricatives support this conclusion. In these experiments the voiced fricatives /ʒ, z, v/, which were presented in isolation and in a CV context, exhibited the recency and suffix effects that had been found earlier for vowels, but not for stop consonants (Crowder, 1973). It is assumed that subjects show significant improvement for recall of the last members of the vowel and voiced fricative series because their echoic memories are more dependable. If this is true, then we would expect subjects to perceive the same stimuli continuously in a discrimination task, because they should be able to refer to echoic memory to help them discriminate stimuli on the basis of differences in the acoustic details of the stimuli.

In conclusion, the results of the experiments in the present study suggest that in addition to cues of long duration, the presence of voicing may be a necessary condition for continuous perception. Since it was found that the voiced fricatives /ɣ, ʕ/, which contain information about the lower formants, were perceived continuously, but that the voiceless fricatives /x, ħ/, which also contain this information, were perceived categorically, it is unclear whether information about the lower formants contributes to continuous perception, as originally hypothesized. It is hoped that future research involving the perception of /ʒ, z/ and whispered vowels will shed some light on this matter.

REFERENCES

ABDEL-MALEK, Z.N. (1972). *The Closed-List Classes of Colloquial Egyptian Arabic* (The Hague).

ABOUL-FETOUH, H.M. (1969). *A Morphological Study of Egyptian Colloquial Arabic* (The Hague).

BARCLAY, J.R. (1972). Noncategorical perception of a voiced stop: A replication. *Perception and Psychophysics*, **11**, 269-273.

CARNEY, A.E., WIDIN, G.P. and VIEMEISTER, N.F. (1977). Non-categorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, **62**, 961-970.

CROWDER, R.G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behavior*, **10**, 587-596.

CROWDER, R.G. (1973). Representation of speech sounds in precategorical acoustic storage. *Journal of Experimental Psychology*, **98**, 14-24.

CROWDER, R.G. and MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, **5**, 365-373.

FRY, D.B., ABRAMSON, A.S., EIMAS, P.D. and LIBERMAN, A.M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, **5**, 171-189.

FUJISAKI, H. (1979). On the modes and mechanisms of speech perception – Analysis and interpretation of categorical effects in discrimination. In B. Lindblom and S. Ohman (eds.), *Frontiers of Speech Communication Research* (London), pp. 177-189.

FUJISAKI, H. and KAWASHIMA, T. (1968). The influence of various factors on the identification and discrimination of synthetic speech sounds. In *Reports of the Sixth International Congress on Acoustics* (Tokyo), Vol. II, pp. 95-98.

FUJISAKI, H. and KAWASHIMA, T. (1969). On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute,* University of Tokyo, **28**, 67-73.

FUJISAKI, H. and KAWASHIMA, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute,* University of Tokyo, **29**, 207-214.

GAIRDNER, W.H.T. (1925). *The Phonetics of Arabic* (London).

GLACE, D.A. (1968). Parallel resonance synthesizer for speech research. *Journal of the Acoustical Society of America,* **44**, 391 (Abstract).

HARRELL, R.S. (1957). *The Phonology of Colloquial Egyptian Arabic* (New York).

KHALAFALLAH, A.A. (1969). *A Descriptive Grammar of SAEIΣDI Egyptian Colloquial Arabic* (The Hague).

LANE, H.L. (1965). The motor theory of speech perception: A critical review. *Psychological Review,* **72**, 275-309.

LAUFER, A. and CONDAX, I.D. (1981). The function of the epiglottis in speech. *Language and Speech,* **24**, 39-62.

LIBERMAN, A.M., HARRIS, K.S., HOFFMAN, H.S. and GRIFFITH, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology,* **54**, 358-368.

MILLER, J.D., WEIR, C.C., PASTORE, R.W., KELLY, W.J. and DOOLING, R.J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. *Journal of the Acoustical Society of America,* **60**, 410-417.

PASTORE, R.E., AHROON, W.A., BAFFUTO, K.J., FRIEDMAN, C., PULEO, J.S. and FINK, E.A. (1977). Common-factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance,* **3**, 686-696.

PISONI, D.B. (1971). On the nature of categorical perception of speech sounds. Ph.D. dissertation, University of Michigan.

PISONI, D.B. (1977). Identification and discrimination of the relative onset of two component tones: Implications for the perception of voicing in stops. *Journal of the Acoustical Society of America,* **61**, 1352-1361.

POLLACK, I. and PISONI, D.B. (1971). On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science,* **24**, 299-300.

REPP, B.H. (In press). Two strategies in fricative discrimination. *Perception and Psychophysics.*

REPP, B.H., HEALY, A.F. and CROWDER, R.G. (1979). Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance,* **5**, 129-145.

SACHS, R.M. (1969). Vowel identification and discrimination in isolation versus word context. *Quarterly Progress Report,* Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 220-229.

SAMUEL, A.G. (1977). The effect of discrimination training on speech perception: Noncategorical perception. *Perception and Psychophysics,* **22**, 321-330.

STEVENS, K.N. (1968). On the relations between speech movements and speech perception. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung,* **213**, 102-106.

STRANGE, W. and JENKINS, J.J. (1978). The role of experience in the perception of speech. In R.D. Walk and H.L. Pick, Jr. (eds.), *Perception and Experience* (New York), pp. 125-169.

STREVENS, P. (1960). Spectra of fricative noise in human speech. *Language and Speech,* **3**, 32-49.