

ON UNIVERSALS IN VOWEL PERCEPTION

Birgit Hutter and Peter Holtse

1. Introduction

It has been shown in a number of studies (e.g. Libermann et al. 1957) that consonants are normally perceived in a categorical way. Thus listeners are generally unable to discriminate much better than they can identify the sounds. The picture is less clear among the vowels. Stevens (1968) reports a tendency towards categorical perception of vowels in words. One recent experiment using isolated vowels (Fujisaki 1971) has shown some correspondence between vowel phoneme boundaries and the ability of listeners to discriminate between vowels. But most perceptual studies on isolated vowels have found a tendency towards continuous perception similar to the way non-speech sounds are perceived. For isolated vowels listeners will discriminate much finer differences in quality than they can identify qualities as phonemes.¹

However, in an interesting study by Stevens, Libermann, Studdert-Kennedy, and Öhman (1969) it is suggested that the perception of vowels is not altogether continuous, but shows peaks and valleys in the discrimination function of a shape similar to the discontinuous discrimination of consonants - although the overall scores are higher than is normally the case with consonants. The peaks and valleys in the discrimination are said to be independent of the linguistic experi-

1) Nobody seems to have investigated how small differences in vowel quality listeners are in fact able to identify. But they are probably considerably smaller than differences between phonemes.

ence of the listeners. And the theory is advanced that certain areas in the vowel continuum are better suited to contain phonemes since the auditory mechanism is less critical about changes in these areas. We should like to offer some comments on the methods used in this study and the conclusions drawn from them.

2. Comments on the article by Stevens et al. (1969)

2.1. Method of posing the problem to the subjects

The experiment reported by Stevens et al. (1969) contained two series of synthetic vowels, one front unrounded (approximately [i-e-ε]) and one narrow, unrounded to rounded (approximately [i-y-u]). A group of Swedish and American listeners were asked to identify the unrounded series with their own front vowel phonemes. Then the Swedes were asked to identify the rounded series with their phonemes /i/, /y/ or /u/ while the Americans were first played a record of the Swedish vowels and then asked to identify the test vowels with the Swedish phonemes. (The fact that they were using numbers instead of phonetic symbols alters nothing in the basic problem.)

This seems rather an unfortunate way of posing the problem. If the Americans did in fact identify the rounded series with the Swedish phonemes they had heard, they must have acquired a Swedish linguistic background, at least as far as the vowel qualities [i-y-u] were concerned. This would mean that the two groups were no longer representatives of different backgrounds, and the object of the experiment: comparing the influence from different linguistic backgrounds, would have been lost. The proper procedure must have been to ask the American listeners to identify the rounded series with their own narrow vowel phonemes - as far as this could be done with the stimuli at hand.

2.2. The discrimination test

2.2.1. Vowel stimuli

The stimuli of the tests were 25 vowels synthesized with approximately equal logarithmic steps. According to the authors "... there were small deviations from uniform spacing. These deviations arose because the formant frequencies for each stimulus could only be set to within a few cps." (p. 4). The magnitudes of these deviations are best judged when the differences in formant frequencies from one stimulus to the next are expressed in per cent. If the differences are equal logarithmic steps they can be expressed as a constant percentage. How far this is the case with the vowel stimuli under consideration may be judged from table I which gives the percentage differences between the formant frequencies of each vowel. As will be seen the deviations are random, but not inconsiderable.

TABLE I

Percentual distances in formant frequencies
between vowel stimuli.²

Based on Stevens et al. (1969), Table 1, p. 3.

Vowel number	F1	F2	F3	Vowel number	F2	F3
1- 2	5.6	1.7	2.0	1-R 2	3.1	3.3
2- 3	4.6	1.6	2.0	R 2-R 3	3.1	3.3
3- 4	5.5	2.1	2.3	R 3-R 4	2.8	4.1
4- 5	6.7	1.7	2.1	R 4-R 5	2.7	3.6
5- 6	5.2	2.0	2.1	R 5-R 6	3.0	4.4
6- 7	5.8	1.6	2.0	R 6-R 7	2.8	3.4
7- 8	6.0	1.8	1.4	R 7-R 8	3.4	3.1
8- 9	5.8	1.6	1.8	R 8-R 9	3.1	2.4
9-10	6.0	1.7	1.0	R 9-R10	3.1	2.4
10-11	6.3	2.0	0.5	R10-R11	2.8	1.9
11-12	6.1	1.5	1.0	R11-R12	2.9	1.7
12-13	5.8	2.1	1.0	R12-R13	2.8	1.1

2.2.2. Results of the discrimination tests

For both the unrounded and rounded vowels Stevens et al. report that the valleys of the discrimination curves correspond roughly to the centres of the phoneme areas as previously established from identification tests, but the correspondence is not particularly good. However, there is surprisingly high agreement between Swedish and American lis-

2) F1-values of the rounded series have not been included in the table since none of the differences exceed 3 Hz.

teners as to where the tops and valleys should be. This is taken as an indication that tops and valleys in the discrimination function are inherent in the perceptual mechanism and not conditioned by linguistic experience. On the contrary, perceptual constraints would favour the placing of vowel phonemes in areas where discrimination is relatively poor.

However, some odd features led us to examine the data more closely. For instance, why did both Swedes and Americans show a pronounced discrimination top between stimuli 5 and 6 in the rounded series? This top is not mentioned very clearly in the article although it is found practically at the top of the Swedish identification function for /y/.

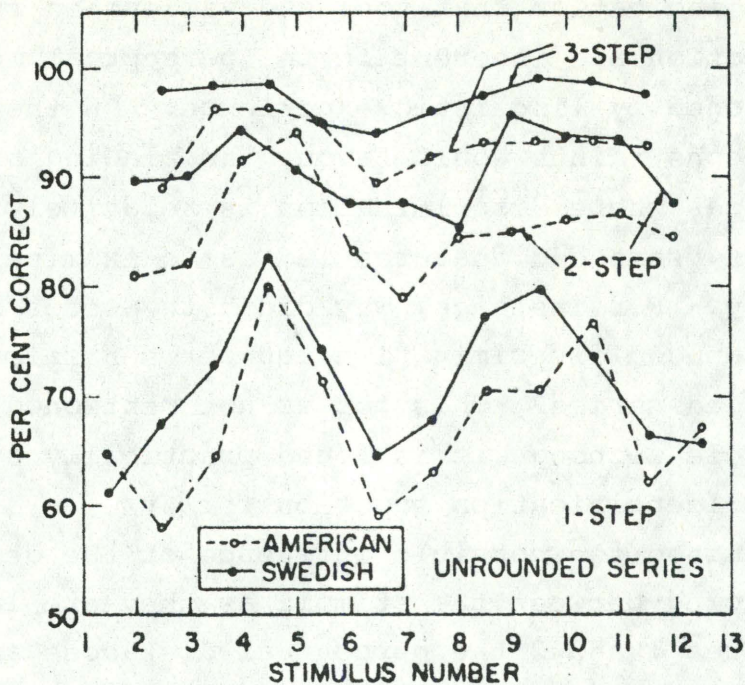
We examined the possible influence of the deviations from uniform spacing between the stimuli as they are listed in table I. In Figures 1 and 2 the percentual distances are compared with the discrimination functions of Stevens et al. (their figs. 6 and 7, p. 10 and 11).

Even to a cursory glance the correspondence between the upper and lower halves of Figures 1 and 2 is quite striking. The top in the discrimination of unrounded vowels between stimuli 4 and 5 corresponds exactly with the large percentual difference between the F1 frequencies of stimuli 4 and 5. And among the rounded vowels the correspondence is even better. The discrimination curves show tops in three places: 3-4, 5-6, and 7-8. The first two tops coincide with large percentual differences in F3, while the third and very small top coincides with a relatively large difference in F2 between stimuli 7 and 8. The relatively poor discrimination between stimuli 8 through 13 may be due to the rather small distances in F3 between these stimuli.

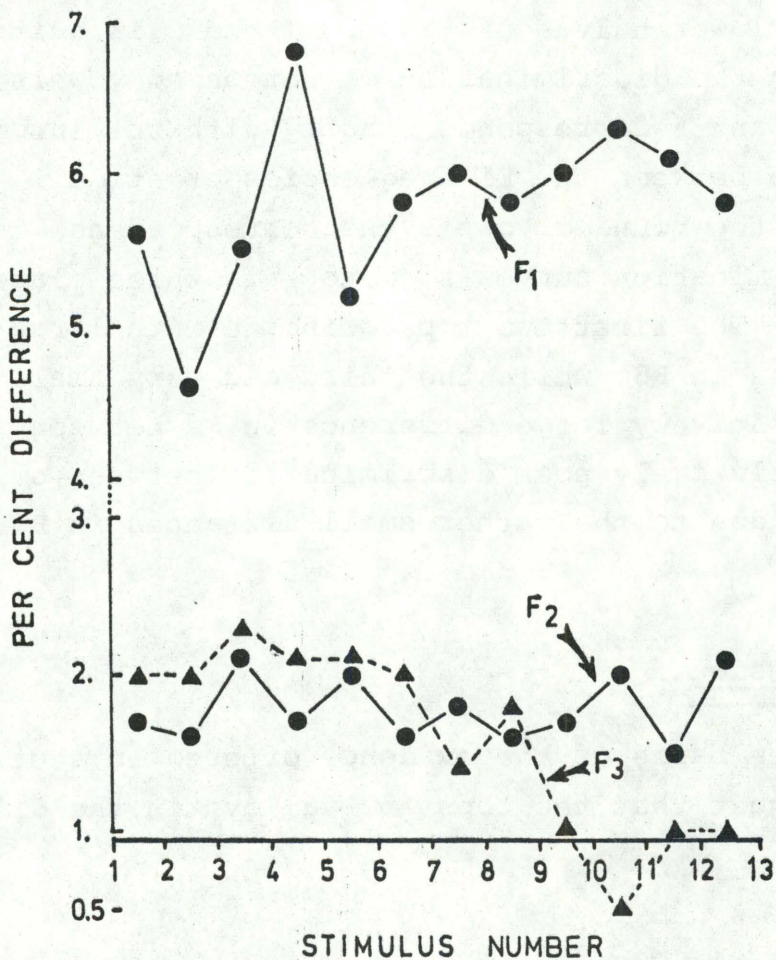
3. Conclusion

On the basis of the evidence offered in section 2.2. we would suggest that the tops and valleys in the discrimination

Fig.1 UNROUNDED VOWELS

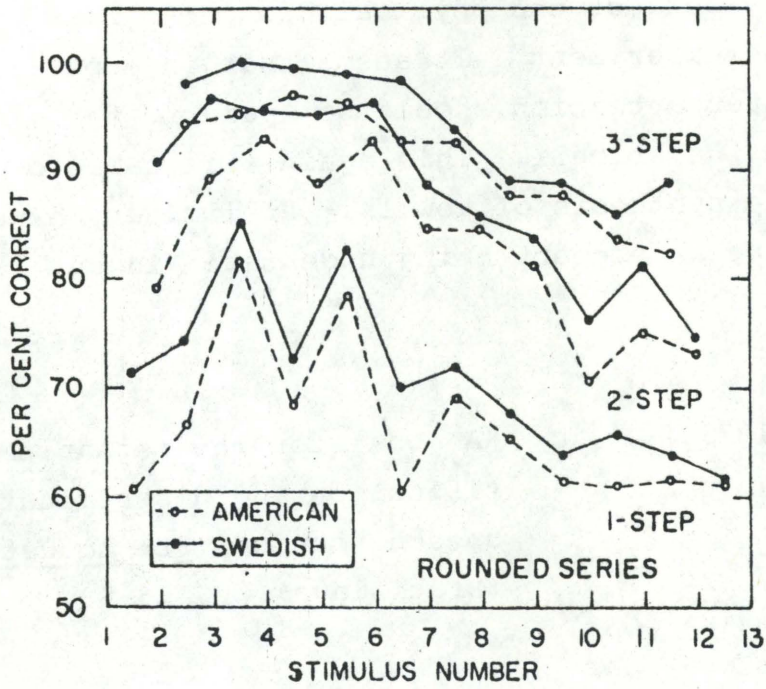


Photocopy of Fig. 6' from Stevens et al. (1969)

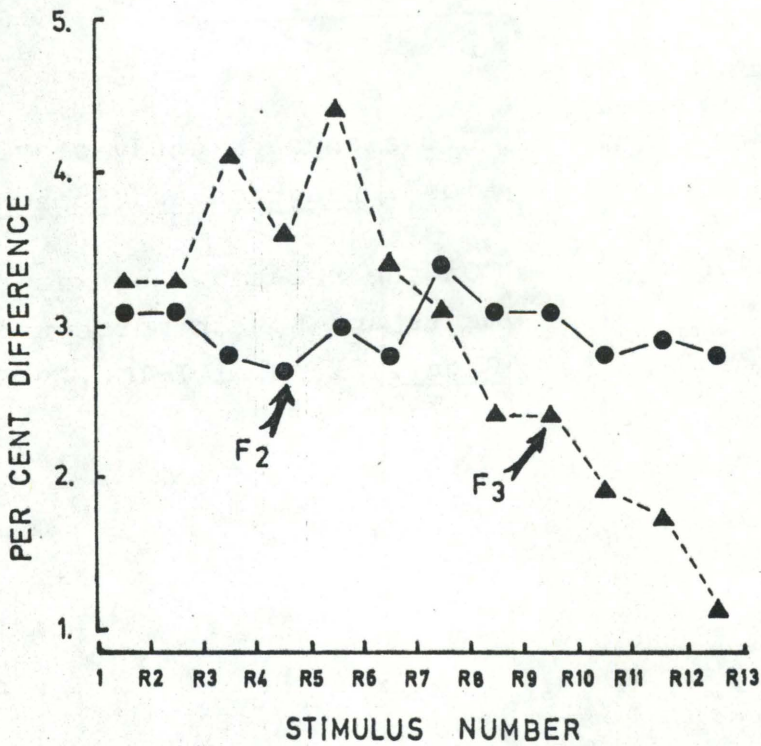


Differences between the vowel stimuli of Stevens et al. (1969) expressed in per cent

Fig. 2 ROUNDED VOWELS



Photocopy of Fig. 7 from Stevens et al. (1969)



Differences between the vowel stimuli of Stevens et al. (1969) expressed in per cent

of vowels as reported by Stevens et al. are not due to any inherent universals in the perceptual mechanism. To some extent they may simply reflect the physical distances between the stimuli used in the experiment. It seems that in experiments of this kind greater attention should be paid to the exact distances between the stimuli. This again will call for greater accuracy in the synthesis of vowels. Preliminary experiments of this kind are at present being undertaken in our laboratory.

References

- Fujisaki, H. 1971: "A model for the mechanisms for identification and discrimination of speech sounds", 9th Acoustics Conf. (Bratislava), p. 56-59.
- Lieberman, A.M., K.S. Harris, A.S. Hoffman, and B.C. Griffith 1957: "The discrimination of speech sounds within and across phoneme boundaries", J. Exp. Psychol., 54, p. 358-368.
- Stevens, K.N., A.M. Liberman, M. Studdert-Kennedy, and S.E.G. Ohman 1969: "Crosslanguage study of vowel perception", Language and Speech, 12, p. 1-23.
- Stevens, K.N. 1968: "On the relations between speech movements and speech perception", Zs. f. Phon., 21, p. 102-106.