STIMULUS CATEGORY, REACTION TIME, AND ORDER EFFECT - AN EXPERIMENT ON PITCH DISCRIMINATION

ANTON BATLINER - LIESELOTTE SCHIEFER

Institut für Phonetik und Sprachliche Kommunikation der Ludwig-Maximilians-Universität München, FRG

ABSTRACT

The "order effect", that causes in a discrimination task the one presentation order to be better discriminated than the reverse order, was tested in the domain of pitch perception with speech and non-speech material as well as with rises and falls. The results showed that (i) rises produce a greater order effect than falls, (ii) non-speech material and rises are better discriminated than speech material and falls, respectively.

INTRODUCTION

The phenomenon of "order effect" (henceforth called OE) has been well known in psychoacoustics since the early thirties, (cf. Stott [7], Zwicker-Feldtkeller [10], Allan-Kristofferson [1]). In the same-different (AX) paradigm, this effect causes the one sequence AB to be discriminated significantly better than the other sequence BA. In psychoacoustic research, this effect has been considered to be an experimental artifact and its influence was eliminated by the following procedure: both orders AB and BA were presented and the mean of the discrimination for both pairs served as criterion for e.g. just noticeable differences, threshold detection etc. cf. [10].
In phonetic research this effect was not dealt with very often (but cf. Repp et al [6], Chuang/Wang [2]). That might be due to the experimental paradigm mostly used in phonetics: in an ABX-task, it cannot show up as clearly as in an AX-task (Repp, 1981 [5]). In our investigations, we used only the AX-paradigm, as it is known [5] that this paradigm is more sensitive than the ABX-paradigm. In several investigations at the Institut für Phonetik in Munich, carried out during the last few years, the OE showed up systematically in studies on speaker recognition (Tillmann/Schiefer/Pompino-Marschall 1984 [9]), tactile discrimination (Tillmann/Piroth 1986 [8]), breathy stops in Hindi (Schiefer, unpublished), German intonation (Batliner, unpublished). In a not yet published paper, we show that the OE is not simply due to the experimental design, and we summarize possible explanations of its origin. In the present paper, we want to address the question of OE from a somewhat different point of view: (i) Does the OE behave differently with speech and non-speech material, i.e. is it a purely psychoacoustic phenomenon, or is there a qualitative difference between speech and non-speech material? (ii) Is there any difference between rises and falls as with regard to the OE?, (iii) What, if any, is the contribution of reaction time to the explanation of the phenomenon? (iv) Is there any difference between the threshold for speech and non-speech material?

MATERIAL

The speech stimulus chosen was 'ja', because the acoustic structure of this stimulus is simple enough so that the factors of interest can be controlled precisely. One of the authors (A.B.) produced several stimuli monotonously in the soundproofed room of the Institute. The stimuli were taped on a Telefunken M15 recorder with a speed of 19 inch per second, digitized on a PDP11/50 with a sample rate of 20 kHz and filtered with a cut off frequency of 8 kHz. For the speech resynthesis of the stimuli a procedure was used where the intensity and the sample points could be defined exactly for each pitch period. The stimulus chosen for the manipulation was segmented into single pitch periods. A logarithmic scale was used for the manipulation of Fo. The stimuli had a constant overall duration of 480 +/-5 ms. The first part containing the fricative, the transition and the first pitch periods of the steady state vowel were left unmanipulated, whereas the remaining pitch periods were subjected to manipulation. Two target stimuli were produced, one falling by one semitone, the other rising by one semitone in its second part. A total of 12teststimuli were derived from the target by increasing the rising contour in six steps of 1/8 tone and decreasing the falling contour analogously in 6 steps of 1/8 tone. These 12 stimuli together with the two target stimuli constituted the body of the speech material. 14 further stimuli were generated, each of which was an exact squarewave analog of the respective speech stimulus.

PROCEDURE

Four different test-tapes were prepared for each of the subgroups (speech-rises, non-speech-rises, speech-falls, non-speech-falls). In the 'same' condition, each stimulus was paired with itself, resulting in 7 combinations. In the 'different' condition, the target stimulus was paired with each of the other stimuli, the order of presentation being AB as well as BA, resulting in 2*6 combinations. Five repetitions of each of the 19 combinations were taped in randomized order, with an interstimulus interval of 500 ms between the members of a pair. Each pair was followed by a pause of 3500 msec; after 10 pairs a pause of 10 seconds followed. The experiments were run in the speech lab of the Institute with a Revox-trainer and headphones, at a comfortable listening level. Subjects were students that were paid for their participation. They were instructed to compare the two members of a pair, to decide as quickly as possible whether they were different or not, and to press the appropriate button on a box forming part of a digital data collecting device. The responses were collected with a PDP11/03 and prepared for statistic analysis.

RESULTS

Figures 1-4 display the different responses for the orders AB and BA; the number of subjects is given in parenthesis (Fig. 1: speech rises (n=14), Fig. 2: speech falls (n=12), Fig. 3: non-speech rises (n=11), Fig. 4: non-speech falls (n=14)). In all graphs the abscissa displays the difference in tone (1/8 to 6/8), and the ordinate the percent different responses. Generally it turned out that the order AB yields more different responses (i.e. is more prominent) than the reverse order BA. This shows up most clearly for speech rises and non-speech falls, less clearly for non-speech-rises. We are at a loss for any convincing explanation for the unsystematic results for the speech falls.

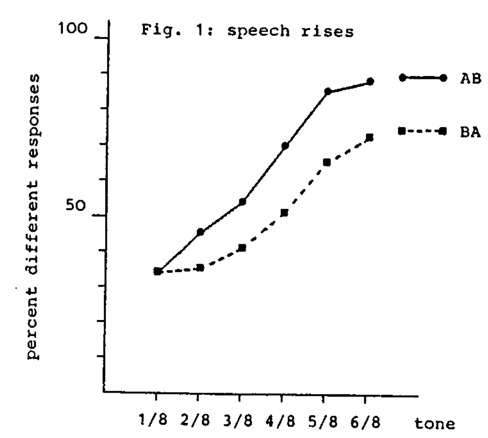

Fig. 1: speech rises

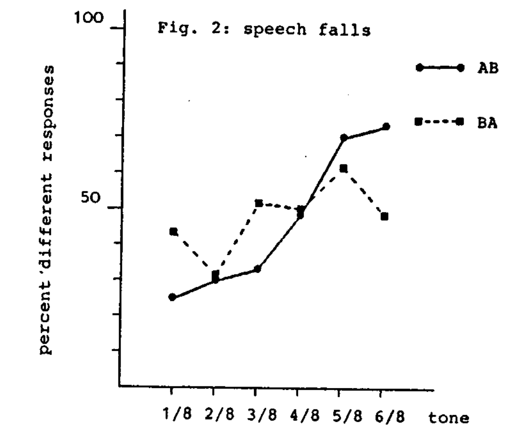

Fig. 2: speech falls

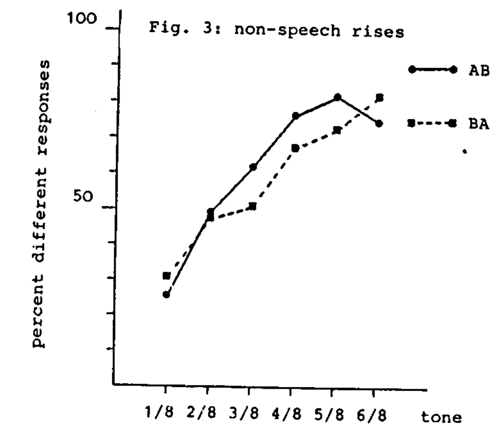

Fig. 3: non-speech rises

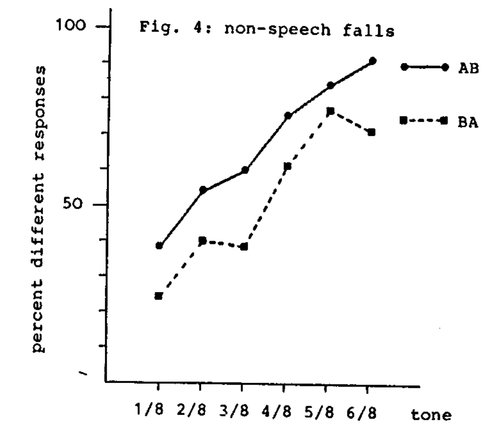

Fig. 4: non-speech falls

A multivariate analysis of variance was applied to the different condition of the four groups together with four factors, two of them being repeated measures (order of presentation AB and BA, difference in tone); the other two, material (speech vs. non-speech) and contour (rise vs. fall) were independent. The level of significance was set to $p < .05$. The necessary assumptions for the multivariate approach were tested with the Cochran and Bartlett tests. Table 1 shows the F-values and level of significance for the effects tested.

<u>Table 1</u>: Statistical results.

	F	p <
BETWEEN-SUBJECTS (df: 1,47)	F	p <
mat. by cont.	1.42	.240
cont.	1.57	.217
mat.	4.22	.046*
ORDER WITHIN SUBJ. (df: 1,47)		
mat. by cont. by ord.	6.95	.011*
cont. by ord.	.03	.860
mat. by ord.	.43	.514
ord.	9.14	.004*
PAIR WITHIN SUBJ. (df: 5,43)		
mat. by cont. by pair	.87	.507
cont. by pair	.60	.694
mat. by pair	1.96	.103
pair	29.09	.001*
ORDER BY PAIR WITHIN SUBJ. (df: 5,43)		
mat. by cont. by ord. by pair	.35	.879
cont. by ord. by pair	2.17	.074
mat. by ord. by pair	1.52	.203
ord. by pair	1.75	.143

Four of the effects tested turned out to be significant: they are asterisked in Table 1: material, material by contour by order, order, and pair. As there was an interacion between material, contour, and order, the significant main effect of order cannot be interpreted. Therefore, Fig. 5 displays the simple main effects for AB and BA; the interaction shows clearly up in the left part of the figure. Given the presentation order BA (right part of the figure), non-speech stimuli yield more different responses than speech stimuli and rises more than falls. This pattern changes for AB (left part) where no difference between speech and non-speech rises can be observed.

Table 2 shows the intersection of the discrimination function of Figs. 1-4 with the 50% line. We can see, that (i) rises, (ii) non-speech material, and (iii) stimuli in presentation order AB can be better discriminated, than falls, speech material, and BA, respectivly.

<u>Table 2</u>: Points of intersection between the discrimination function and the 50% line.

	speech rises	speech falls	non-speech rises	non-speech falls
AB	2.5	4.08	2.07	1.73
BA	3.86	2.91	2.75	3.5

Fig. 6 displays the mean reaction time (RT) for all four groups taken together. The ordinate shows the RT in ms, the abscissa the 'same/different' responses for the two orders AB and BA. It is obvious that responses to the order AB require longer RTs than those to the order BA, and RTs are shorter for 'different' than for 'same' responses, i.e., hits require less RT than false alarms. (In the 'same' response condition, the difference between the orders AB and BA turned out to be significant, $F(1,1303) = 8.89$, $p < .01$.) These results are comparable to those from the identical pairs, where 'same' responses (i.e. hits) have shorter RTs than 'different' ones.

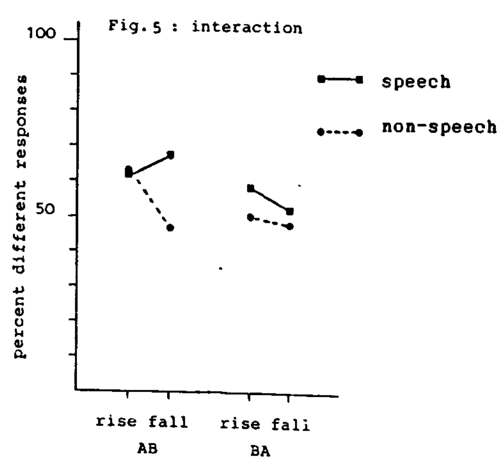

Fig.5 : interaction

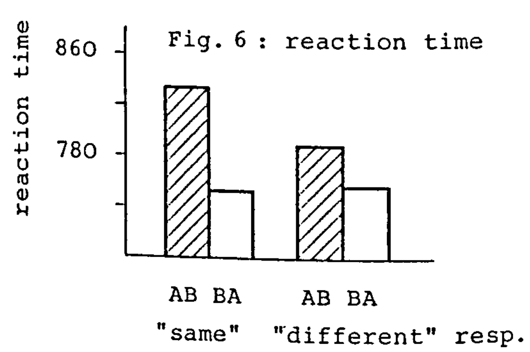

Fig. 6 : reaction time

DISCUSSION

As for material and contour, our results are in agreement with the findings of Klatt [4] and t'Hart [3], who showed that rises are better discriminated than falls and nonspeech better than speech material. The OE turned out to be no purely psychoacoustic phenomenon, as it could be found with the speech and the non-speech material. The present results confirm further our hypothesis, based on earlier findings, that the order AB is better discriminated than the reverse order BA, i.e., stimuli are better discriminated if the stimulus with the greater change in Fo comes last. It doesn't seem to be the height of the offset that is responsable, but the amount of Fo-movement, because otherwise the OE for the falls would favor the order BA and not AB. In the above mentioned paper we will deal with the origin of the OE in detail.

REFERENCES

[1] Allan, L.G. - Kristofferson, A.B.: Psychophysical theories of duration discrimination. Perception & Psychophysics 16: 26-34 (1974)

[2] Chuang, Ch.-K. - Wang, W.S.-Y.: Time-order-error in judgement of prosodic features: pitch, loudness, and duration. Acoust. Soc. Am. 62(S): S48 (1978)

[3] 't Hart, J.: Differential sensitivity to pitch distance, particularly in speech. J. Acoust. Soc. Am. 69: 811-821 (1981)

[4] Klatt, D.H.: Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. J. Acoust. Soc. Am. 53: 8-16 (1973)

[5] Repp, B.H.: Categorical perception: issues, methods, findings. Haskins Laboratories Status Report on Speech Research SR-70: 99-181 (1982)

[6] Repp, B.H. - Healy, A.F. - Crowder, R.G.: Categories and context in the perception of isolated steady-state vowels. Journal of Experimental Psychology: Human Perception and Performance 5: 129-145 (1979)

[7] Stott, L.H.: Time-order-error in the discrimination of short tonal durations. Journal of Experimental Psychology 18: 741-766 (1935)

[8] Tillmann, H.G. - Piroth, H.G.: An order effect in the discriminability of pulse train sequences. J. Acoust. Soc. Am. 79: S73 (1986)

[9] Tillmann, H.G. - Schiefer, L. - Pompino-Marschall, B.: Categorical perception of speaker identity. Proc. 10th Int. Congr. Phon. Sci., Utrecht, pp. 443-448 (Cinnamison, Dordrecht 1984)

[10] Zwicker, E. - Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. Stuttgart 1967