Institut
für
Phonetik und Sprachliche Kommunikation
der Universität München

FORSCHUNGSBERICHTE

29

1991

ISSN 0342 - 782X

Forschungsberichte des
Instituts für Phonetik und
Sprachliche Kommunikation
der Universität München
(FIPKM) 29 (1991) 125-180.

A ramble round the order effect

LIESELOTTE SCHIEFER & ANTON BATLINER

Abstract

In a discrimination task, the order effect, which causes one presentation order to be better discriminated ("prominent order") than the reverse order, was tested in the domain of pitch perception. The following questions were addressed: (i) Can the order effect be influenced by the experimental design? (ii) What causes a specific order to be the prominent one? (iii) Can we call one of the stimuli in a pair the "prominent stimulus"? (iv) Is the order effect just an experimental phenomenon or can it be found in real life as well? (v) Is there any difference between speech and non-speech material concerning the order effect?

CONTENTS

0. PROLOGUE

In this paper, we will give a summary of our work on the order effect (henceforth called OE) in pitch perception. A shorter and revised version with the title "The order effect in pitch discrimination - a speech or a non-speech phenomenon?" is in preparation and will be submitted to "Language and Speech".

1. INTRODUCTION

1.1 THE ORDER EFFECT IN PSYCHOPHYSICS

The phenomenon of OE dates back to Fechner, who established this effect while investigating the psychophysics of weight perception. He introduced and exemplified the effect, under the name of "konstanter Fehler" ("constant error"), as follows:

> Insofern die Verhältnisse einer bestimmten, für die verschiedenen verglichenen Grössen verschiedenen, Zeit- und Raumlage constant durch eine Versuchsreihe bleiben, begründen sie im erlangten Masse das, was man im Allgemeinen einen c o n s t a n t e n F e h l e r nennen kann. (Fechner 1860 [1964]: 90) So ist in meinen Gewichtsversuchen bei hinreichend schwerem Hauptgewichte das erstaufgehobene, also in Betreff der Zeitlage vorangehende, Gewicht, abgesehen von Zufälligkeiten, stets als das leichtere erschienen, wenn schon das Mehrgewicht D bei demselben war, solange dieses Mehrgewicht nicht über eine gewisse Gränze [sic] stieg. (Fechner 1860 [1964]: 124)

A literal translation of this 19th century scientific prose might be more disturbing than clarifying. Instead, we want to express Fechner's example in our own wording: If subjects were to lift up two weights one after the other, the heaviness of the second weight would be overestimated in comparison to the first one. This holds true as long as the weights do not fall below a certain limit. In more general terms, if we were to evaluate two sensations, which are caused by two similar events A and B that follow each other on the time axis, the order AB yields a different result than the order BA.

Following Fechner, the OE (in the English literature called "time order error") was investigated with different material and under a variety of experimental conditions by, e.g., Köhler (1923), Lauenstein (1932), Stott (1933, 1935), Needham (1934a, 1934b, 1935), Woodrow and Stott (1936), Peak (1939, 1940a, 1940b), Allan and Kristofferson (1974). Although the OE turned out to be a stable phenomenon, no straightforward explanation could be given for the regularities that hold for the OE in general or for a certain subset of material and experimental conditions in particular. In psychoacoustic research, the "same-different" paradigm (AX Task) has been mostly used; the OE has normally been considered to be an experimental artifact, and its influence was eliminated by taking the mean of the judgment scores for the orders AB and BA (cf. Zwicker and Feldtkeller 1967).

1.2 THE ORDER EFFECT IN PHONETICS

In phonetic research, the OE has not been dealt with very often (but cf. Jamieson and Petrusic 1975, Smith 1976, Chuang and Wang 1978, Repp, Healy and Crowder 1979, Repp and Crowder 1990). This might be due to the experimental design mostly used in phonetics and especially in the paradigm of categorical perception - the ABX task. In this task, A and B are different, and subjects are asked to decide whether X is identical with A or with B; the subjects must therefore store two stimuli in memory and possibly categorize them as well before the third stimulus can be compared with them. In an AX task, only stimulus A must be held in memory and compared with stimulus X. The AX task is therefore more sensitive than the ABX task, as memory load is less important; i.e., the probability is increased that the perception of the listener is tested and not his/her classification (cf. Repp 1984:266).

We will now, once again, characterize the OE but this time by means of a constructed example that exhibits the characteristic features of the experiments which will be reported in this paper. Let us suppose that we generate a continuum with six physically equidistant stimuli. If these stimuli are, for example, level tones, then their fundamental frequency (Fo) increases in equal steps across the continuum. In an AX task, three different types of stimulus pairs are presented: AA (or BB), i.e., pairs of identical stimuli; and AB and BA, i.e., pairs of different stimuli in each of two possible orders (cf. Table 1). Let us suppose that the step-size is one, so only adjoining stimuli are presented in the orders AB and BA.

Table 1: Stimuli and combinations in an AX-task with step-size one

Stimuli:	1	2	3	4	5	6
same:	11	22	33	44	55	66
different (AB):	12	23	34	45	56	
different (BA):	21	32	43	54	65	

The subjects must judge whether the stimuli in each AX pair are the same or different. An OE can be observed if the "different" judgments for the order AB are significantly more (or less) accurate than for the order BA. We will call the order that is discriminated better the **"prominent" order**. The stimulus that comes second in this order will be called the **"prominent" stimulus**. (For the moment, this is just a convenient way to refer to different stimuli; an explanation will be given later.)

1.3 AIM OF THIS STUDY

In this paper we wish to address the following questions:

(i) What causes a specific order to be a prominent one?

(ii) Can the OE be influenced - or even reversed - by changing certain features of the experimental design?

(iii) Can the OE and thereby the prominence of a stimulus in our material be traced back to general psychophysical factors, or is it a special speech phenomenon (or something in between)?

(iv) Is the OE an experimental artifact that simply has to be controlled or is it a systematic factor that has to be taken into account? That is, is the OE a laboratory phenomenon, or can it be found in real life as well?

2. METHODS

2.1 MATERIALS

It was shown by Studdert-Kennedy and Hadding (1973) that even in a 'simple' one-word sentence, such as *November*, quite complex interactions between Fo onset, Fo peak, and Fo offset can take place with regard to the perceptual distinction between questions and non-questions. Therefore, we decided in favor of a less complex stimulus - German *ja (yes)*. For investigations concerning categorical perception, this stimulus might be problematic because of its many linguistic and especially paralinguistic functions (cf. Willkop 1988:86ff). In our case, however, this does not matter, as our primary interest is not the peak of the discrimination function, but the different discriminability of the two presentation orders. Nevertheless, the two identification experiments that we conducted (cf. Figures 3 and 7) indeed showed a change of the labeling function that converged with the peak of the discrimination function.

One of the authors (A.B.) produced several stimuli monotonously in a soundproof room of the Institute of Phonetics in Munich. The stimuli were taped on a Telefunken M15 recorder with a speed of 19 cm per second, digitized on a PDP11/50 with a sample rate of 20 kHz and filtered with a cutoff frequency of 8 kHz. For the speech resynthesis of the stimuli, a procedure was used where the intensity and the Fo value 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 (semitone = 17.31*ln(Hz)). The first part of the stimuli (cf. Figure 1) containing the approximant, the transition, and the first pitch periods of the steady state vowel were left intact, whereas the remaining pitch periods were subjected to different manipulations of the Fo contour. The intensity of the whole stimulus was left unchanged. We will not give the exact overall duration values for the different stimuli, but the approximate value. The exact values differ by maximally +/- 5 ms because of the pitch-synchronous manipulation; these differences are below the threshold of detection.

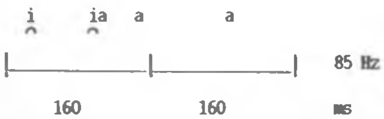

Figure 1: Segmental and durational structure of the stimuli

2.1.1 "Rise-Fall" Continuum

The manipulated parts of the first continuum generated are shown in Figure 2a (1 to 9). The continuum consists of nine stimuli, three falls, one level and five rises. Two clearly distinct linguistic categories exist in this continuum: The clear falls denote assertions, the clear rises questions. For all stimuli the overall duration and the duration of the manipulated part were kept constant, while Fo offset and Fo slope were different. The step from one offset height to the next was one semitone. The slope could not be independently controlled; it changed as a result of the offset manipulation.

This continuum was used in only one experiment because we wanted to isolate the contributions of Fo offset, duration of Fo contour, and Fo slope. Therefore, additional continua had to be generated. In order to keep the number of stimuli within reasonable limits, falls, and thereby one linguistic category, were excluded in most of the the other experimental series. From this point on, if we feel that the context is clear enough, we will sometimes simply speak of "duration" in the sense of "duration of Fo contour", and of "Fo offset" in the sense of "height of Fo offset".

Three further continua were generated, in which duration, Fo offset, and Fo slope, respectively, were held constant, while the other two parameters varied. A total separation of the three factors was not possible, as two of them always covary automatically. Note that the names given to these continua denote the factors that have been manipulated. The stimuli at the endpoints of all the three continua, number 1 (level) and number 6 (114 Hz offset) are identical, only stimuli 2 to 5 differ.

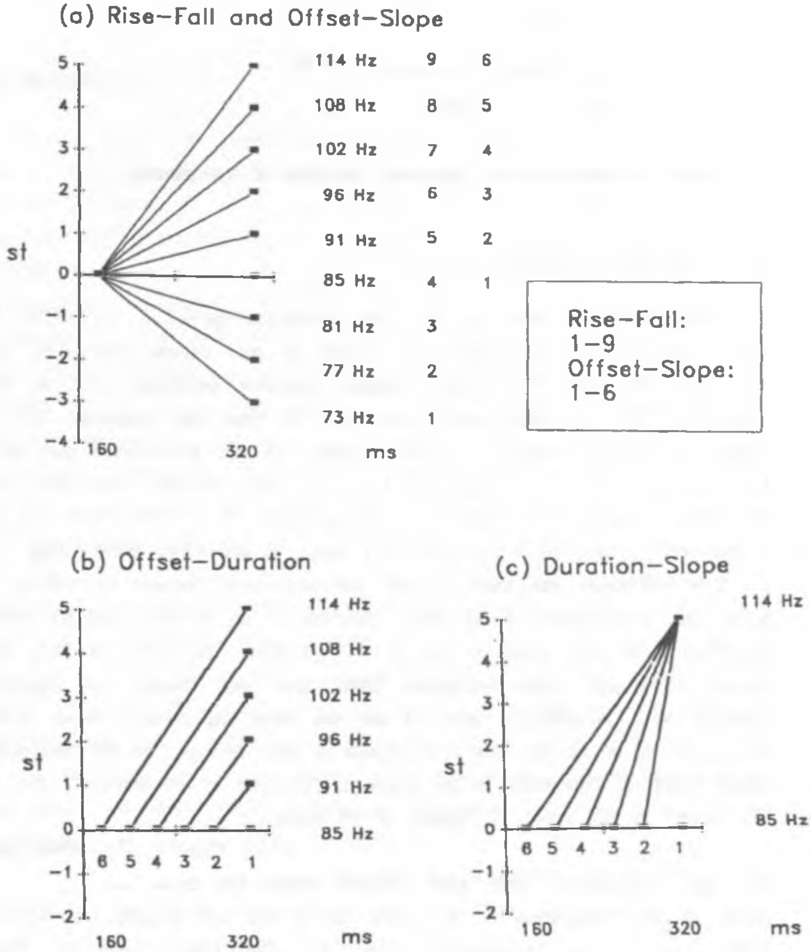

Figure 2: Continua

2.1.2 "Offset-Slope" Continuum

This continuum includes six stimuli (1 to 6 in Figure 2a), that correspond to stimuli 4 to 9 in the Rise-Fall Continuum: the level stimulus 4 in Figure 2a, e.g., corresponds to the stimulus 1 in the Offset-Slope Continuum, etc.

2.1.3 "Offset-Duration" Continuum

The slope chosen for this continuum (Figure 2b) was the same as that of stimulus 6 in the Offset-Slope Continuum. The height of the Fo offsets corresponded to the offsets of the Offset-Slope Continuum, i.e., they differed by one semitone each. As the overall duration of the whole stimulus was kept constant, the decrease of the duration of the manipulated portion lead to a corresponding increase of the unmanipulated portion. The difference between the stimuli was 30 ms for stimuli 2 to 6; the duration of the manipulated part of stimulus 2 was 40 ms.

2.1.4 "Duration-Slope" Continuum

For all stimuli, the offset was held constant at 114 Hz corresponding to stimulus 6 in the other two continua; the duration of the manipulated portions corresponded to the manipulated portions of the Offset-Duration Continuum. The continuum is shown in Figure 2c.

2.2 PROCEDURE

In this part, we would like to describe the default procedure for most of the experiments; exceptions are specified in the sections "material and procedure".

Subjects were partly students participating in a course in phonetics and partly students that were paid. The experiments were run with an interval of at least one week but quite often separated by much longer intervals. Although most of the subjects took part in more than one experiment, it seems legitimate to consider the experiments independent.

For all of the experiments, the AX task was used. Normally, five repetitions of each order of each "different" pair (AB, BA) and of each "same" pair were presented in randomized order with an interstimulus interval of 500 ms between the members of a pair. The pairs were separated by pauses of 3500 ms; after 10 pairs, a pause of 10 sec followed. The experiments were run either in a quiet room using a Revox 77 taperecorder at a comfortable listening level, or in the speech laboratory of the Institute with a Revox trainer. The subjects were instructed to compare the two members of a pair and to decide whether they were identical ("same"), or different in any respect. They were asked to press the appropriate button on a box forming part of a digital data collecting device. The responses were collected with a PDP11/3 and prepared for statistical analysis.

There were two limitations of the experimental design:

(i) To avoid undue length of the sessions (and thereby the risk of increasing inattention), the number of repetitions had to be decreased and/or the step-size had to be increased in some cases.

(ii) The available software for the generation of the test tapes did not allow every possible combination of the stimuli. This limitation applies to the experiments described in part 3.2.2.

With the exception of the experiments described in parts 3.2.2. and 3.4., a two-way analysis of variance was used, the first factor being the OE, i.e., the order of the pairs AB or BA, the second the pairs. Generally, a significance level of 5 percent was chosen. The necessary conditions for an additional statistical evaluation of comparisons between experiments was rather seldom met; these comparisons are therefore carried out on an interpretative basis only (again, with the exception of the experiments described in part 3.4.). Note that our main interest is to show that the OE is consistent irrespective of the changes in the experimental design, etc. (comparison within experiments); our main interest is not to show that some changes in the experimental design have a statistically significant influence on the results (comparison between experiments). An abundance of significant comparisons (hopefully even "significant at the 0.001 level") might look attractive, but experience has shown that safe reasoning is the better way (cf. Guttman 1977:passim).

Normally, the figures show the discrimination functions for the order AB (solid line with full circle), BA (solid line with open square), and for the same pairs (dotted line with "x"). AB means always stimulus *n* followed by stimulus *n + step-size*, i.e., it is consistent with the enumeration of the stimuli on Figures 2 and 6. If an identification test was run, its function is displayed by a dashed line.

An analysis of the single subjects was carried out for most of the experiments as well. It sometimes showed enormous differences between the subjects, but we will not report on this in order to save space. It might suffice that the OE is stable for the whole samples across the subjects.

3. EXPERIMENTS

3.1 STARTING POINT: THE POTBELLY PHENOMENON

3.1.1 Materials and procedure

In this experiment, the Rise-Fall Continuum (cf. Figure 2a) and the default procedure were used. 12 subjects took part. An identification test was run in addition, i.e., a forced choice task with the categories "question" and "assertion". 13 other subjects took part; each stimulus was presented six times in randomized order.

3.1.2 Results and discussion

The resulting functions are given in Figure 3. The OE main effect, the stimulus pairs effect, and the interaction are significant; OE: $F(1,11) = 60.67$, $p < .001$; stimulus pairs effect: $F(7,11) = 6.16$, $p < .01$; interaction: $F(7,11) = 3.32$, $p < .05$.

With this potbelly shape function, a clear OE could be found; the order AB was discriminated better than the order BA; the prominent order shows a higher Fo offset in the second member of the pair.

The false alarms for the "same" pairs produced a peak in the region of stimulus 6, i.e., near the peak of the discrimination function of the "different" pairs and near the category boundary of the identification function. A possible explanation for this phenomenon could be that the subjects compared not only the acoustic structure of the two members of a pair, but they also classified them into one of the two categories. In our case, the categories are "assertion" vs. "question". The first stimulus is classified, the second one is checked as to whether it belongs to the same category as the first one. If this is the case, the answer is "same", otherwise it is "different". If the first stimulus is a borderline case between the two categories, then the first stimulus is still classified as a member of one category. (This is the crucial point: The membership to a category is more salient then the exact acoustic structure.) The second stimulus is then checked as to whether it belongs unequivocally to the same category. However, this is not the case, and the answer is therefore more often "different" than in the proper region of the one or the other category (cf. Schiefer and Batliner 1988:279f).

The identification function is given as a dashed line in Figure 3. There is no ideal steep slope, and there is some "noise" within categories, but the category boundary between stimuli 5 and 6 converges with the peak of the discrimination function. Note that this boundary is not at the level stimulus 4 but at a stimulus two semitones above level; this result is consistent with other experiments on German intonation, where short sentences were used, cf. Batliner (1989a, 1989c).

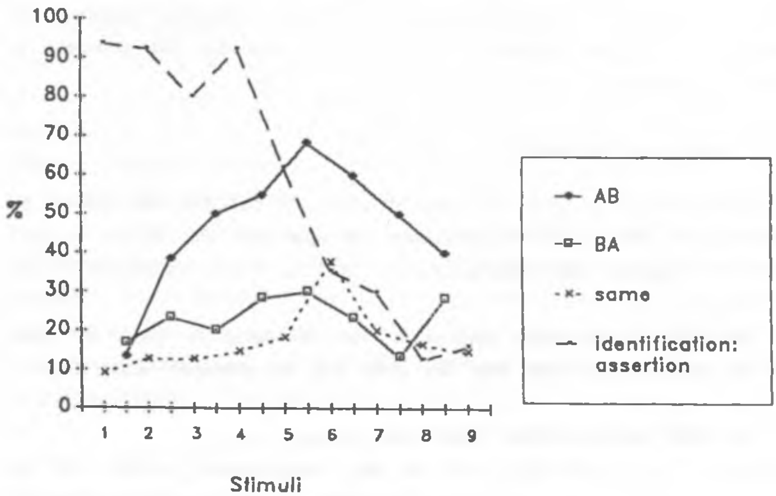

Figure 3: Rise-Fall Continuum, one step, discrimination functions

3.1.3 Guide through the experiments

The potbelly phenomenon can be a reasonable starting point for at least the following topics, that will be dealt with in this paper:

(i) **Triggering factors:** The relevance of the factors that contribute to the OE cannot be decided upon with one experiment. In part 3.2., we will therefore report the results of experiments where these factors were varied as systematically as possible.

(ii) **Design effect:** The OE is historically – and up till now – an experimental phenomenon, i.e., it originated and has grown up in the laboratories. It is always connected with a special experimental design. Therefore, it might be reasonable to investigate the possibility of whether the OE is not simply the outcome of special designs. This question will be addressed in part 3.3.

(iii) **Speech vs. non-speech mode of perception:** In the history of experimentation, the OE was not originally considered to be a speech phenomenon – rather the other way round: It was considered to be a psychophysic/psychoacoustic phenomenon. In parts 3.4, 4. and 5., we wish to investigate whether our findings are compatible with a strict psychoacoustic point of view, or whether there are results that can better be explained with a speech mode of perception. (We do not want to imply that in this case the speech mode is necessarily special and that some sort of modularity exists as, e.g., Liberman and Mattingly 1985 assume.)

(iv) **Prominent order/prominent stimulus:** It is clearly not enough to simply establish an OE without knowing its underlying reasons. In part 4., we therefore give an account of the relevant literature and try to find out the reasons for the OE.

(v) **OE and real life:** The experimental procedure used in the investigation of the OE was normally not a means to simulate real life phenomena. In part 5, we will deal with a special speech phenomenon (perception of accents) where striking parallels to the OE can be found.

3.2 TRIGGERING FACTORS

With these experiments, we wanted to investigate which factors are relevant for the OE, as far as our relatively simple stimuli are concerned. All three continua were tested.

3.2.1 Factors isolated (one-step)

3.2.1.1 Materials and procedure

The Offset-Slope Continuum was tested with 14 subjects, the Offset-Duration Continuum with 15, and the Duration-Slope Continuum with 13 subjects; the default procedure was used.

3.2.1.2 Results and discussion

As Figure 4a shows, the OE main effect in the Offset-Slope Continuum is comparable to the one in Figure 3: $F(1,13) = 26.21$, $p < .001$; neither the stimulus pair effect ($F(4,13) = 1.35$, n.s.) nor the interaction ($F(4,13) = 1.84$, n.s.) are significant. The order AB (the stimulus with the higher Fo offset comes second) is better discriminated than the order BA. Both orders are discriminated better than in the Rise-Fall experiment; for the "same" pairs, there are more false alarms, and there is no peak, but a plateau between stimulus 3 and 6 in the region of clear rises. The reason might be that in this continuum there are not two clearcut categories (assertion/question), but only one clearcut (question) and one subcategory (unfinished, "progredient" assertion). Again, the OE has a potbelly shape function. It is more pronounced in the middle of the continuum than at the edges.

The results for the Offset-Duration Continuum are given in Figure 4b. The OE main effect ($F(1,14) = 6.67$, $p < .05$) and the stimulus pair effect ($F(4,14) = 36.92$, $p < .001$) are significant, whereas the interaction is not ($F(4,14) = 1.59$, n.s.). With the exception of 1/2 and 2/3, the order AB (the stimulus with longer duration of the contour and higher Fo offset comes second) is better discriminated than the order BA. The longer the duration of the contour and the higher the offset, the greater is the OE. It is likely that for pairs 1/2 and 2/3, the contours are not perceptually clear enough. An OE can be observed only if the duration of the manipulated part is equal to or greater than 70 ms, and the height of the Fo offset is equal to or greater than 2 semitones above level. The false alarms for the "same" pairs show a peak at stimulus 4 (contour of 100 ms duration, offset 3 semitones above level). Here, as well as for the "same" pairs in the Rise-Fall experiment, the above mentioned explanation applies. The category boundary between "assertion" and "question" in German is not near level, but in the region of an offset that is two to four semitones above level, cf. Batliner (1989a:117ff, 1989c:99ff).

The functions for the Duration-Slope Continuum are given in Figure 4c. The OE main effect ($F(1,12) = 12.1$, $p < .01$) and the stimulus pair effect ($F(4,12) = 17.95$, $p < .001$) are significant, the interaction is not ($F(4.12) = 2.59$, n.s.). The order AB is prominent (the stimulus with the longer duration of the contour comes second); however, the OE breaks down at pairs 1/2 and 4/5. In the case of 1/2, the reason might be that stimulus 2 with only 40 ms contour is perceived not as a contour stimulus, but as a level stimulus, perhaps with some disturbance at the end. (Some of the subjects remarked that no contour was heard at stimulus 2, but only level with some slight "curling" at the end). Note that 1/2 gets almost the same scores as the "same" pair 2/2. At 2/3, the OE is most prominent. This might be due to the comparison of a stimulus that is perceptually "level" (stimulus 2) with a stimulus that is a pronounced "contour" (number 3). The comparison of level with contour is a special case, cf Lehiste (1976). However, the OE cannot be limited to that particular case, as the pair 5/6, i.e., two clear contour stimuli, produced an OE as well. We have no explanation for the breakdown of the OE at 4/5. (Note that in the experiment described in 3.3.2., a comparable breakdown cannot be observed.) The peak of the discrimination function is at 3/4, as is the case for the Offset-Duration Continuum. Possibly a psychoacoustic threshold is involved, cf. Hombert, Ohala and Ewan (1979), who report that only Fo contours above 60 ms length could be perceived as contours.

The "same" pairs end in a plateau without a pronounced peak and are comparable to the "same" pairs in the Offset-Slope Continuum.

3.2.2 Mixed factors

3.2.2.1 Materials and procedure

In the experiments discussed up till now, the factors Fo offset, duration of the Fo contour, and Fo slope were separated, and only members of one continuum were tested together. In these tests, one acoustic factor was kept constant and two factors were manipulated; in the Offset-Slope Continuum, e.g., the duration of the contour was kept constant whereas the Fo offset and Fo slope were manipulated.

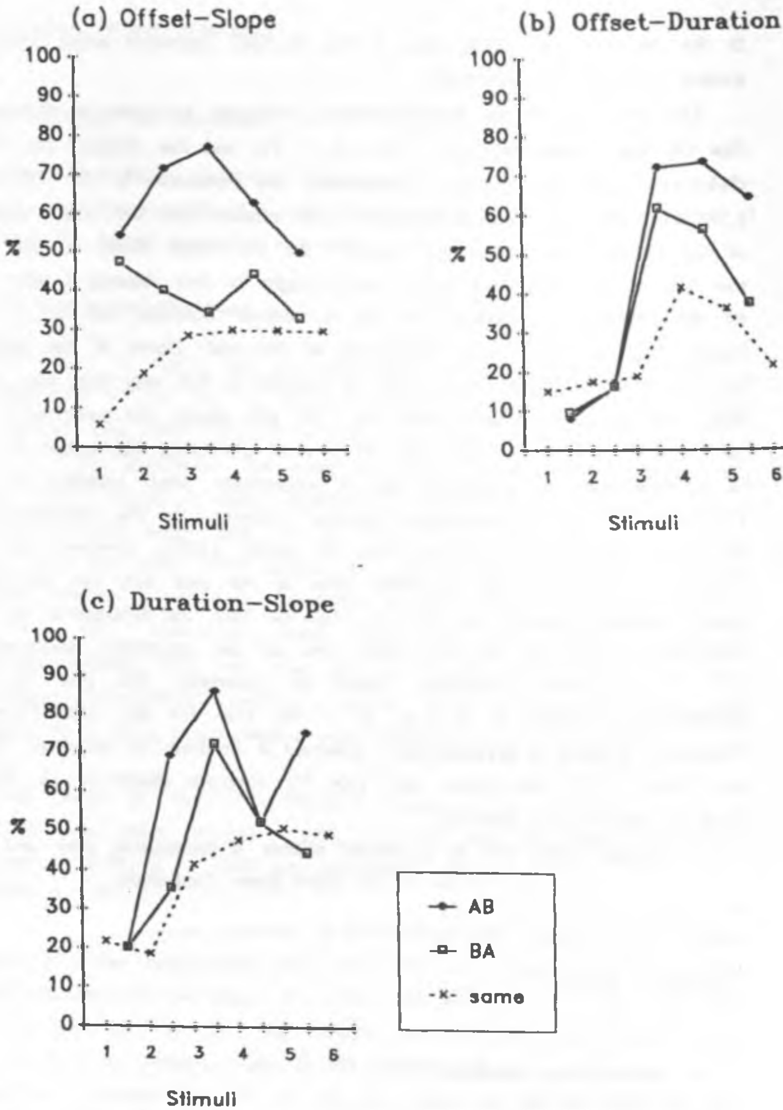

Figure 4: One step experiments, discrimination functions

Additionally, the step-size was the same throughout the continua, i.e., in all pairs the acoustic difference was the same: A difference of 30 ms in the Duration-Slope Continuum and Offset-Duration Continuum (except for pair 1/2, where the difference was 40 ms) and a difference of 1 semitone in the Offset-Slope Continuum and Offset-Duration Continuum. With the following experiments, we attempted to use our stimuli in order to investigate possible trading relations and thereby decide upon the relevance of the factors duration of Fo contour, Fo slope, and Fo offset.

Three experiments were conducted. In experiment 1, stimuli from the Offset-Slope Continuum and Offset-Duration Continuum were used. In the second experiment, Offset-Slope stimuli and Duration-Slope stimuli were mixed, and in experiment 3, the stimuli were taken from the Offset-Duration and Duration-Slope Continua.

In the following, the stimulus combinations are given with the first characters of the names and the number of the respective stimulus; OS_3DS_4, e.g., means stimulus 3 from the Offset-Slope Continuum combined with stimulus 4 from the Duration-Slope Continuum. The stimulus combinations are given in Table 2. These combinations met certain conditions:

(i) OD_nOS_n: Stimuli having the same Fo offset, which is by one or two semitones lower than that of the Duration-Slope Continuum.

(ii) $OD_{n+1}OS_n$: Stimuli with different Fo offsets (the difference being one semitone) and different durations of the contour (30 ms to 90 ms). In the stimuli with a longer duration of the Fo contour, the Fo offset was reduced by one semitone. With this manipulation, we hope to compensate for the longer duration of stimulus OS_n.

(iii) OD_nDS_n: Stimuli with the same duration of the contour but Fo offsets that differ from one to four semitones.

(iv) $OD_{n+1}DS_n$: Stimuli with Fo offsets that differ from one to three semitones. Here, the stimulus with the higher Fo offset was 30 ms shorter than the other stimulus of the pair. Again, we wanted to compensate for the greater Fo differences through a longer duration of the Fo contour in the stimulus OD_{n+1}.

(v) OS_nDS_n: Stimuli with different Fo offsets and different durations. Whereas the duration of the stimulus OS_n was by 30 to 120 ms longer, the Fo offset of the stimulus DS_n was by 1 to 4 semitones higher; i.e., in this case, we can expect clear trading relations between the factors duration of the Fo contour and Fo offset.

(vi) OS_nDS_{n+1}: Stimuli with different Fo offsets and durations of the Fo contour. Here the compensation of differences in the Fo offset through longer durations of the stimulus OS_n is less (30 to 90 ms) compared with the combination OS_nDS_n.

Table 2: Mixed factors

Experiment 1: Offset-Slope and Offset-Duration Continuum

Comb.	Pair A B	st-diff	ms-diff	perc. diff AB	perc. diff BA	OE
OD_nOS_n	OD_2OS_2	–	OS+120	12	17	OS –5
	OD_3OS_3	–	OS +90	65	53	OS+12
	*OD_4OS_4	–	OS +60	88	58	OS+30
	*OD_5OS_5	–	OS +30	30	12	OS+18
OD_{n+1}- OS_n	OD_3OS_2	OS-1	OS +90	7	17	OS-10
	*OD_4OS_3	OS-1	OS +60	42	40	OS +2
	*OD_5OS_4	OS-1	OS +30	37	28	OS +9

Experiment 2: Offset-Slope and Duration-Slope Continuum

Comb.	Pair A B	st-diff	ms-diff	perc. diff AB	perc. diff BA	OE
OS_nDS_n	OS_2DS_2	DS+4	DS-120	42	50	DS -8
	OS_3DS_3	DS+3	DS -90	47	78	DS-31
	*OS_4DS_4	DS+2	DS -60	57	48	DS +9
	*OS_5DS_5	DS+1	DS -30	28	37	DS -9
OS_n- DS_{n+1}	OS_2DS_3	DS+4	DS -90	43	55	DS-12
	*OS_3DS_4	DS+3	DS -60	70	43	DS+27
	*OS_4DS_5	DS+2	DS -30	43	30	DS+13

Experiment 3: Offset-Duration and Duration-Slope Continuum

Comb.	Pair A B	st-diff	ms-diff	perc. diff AB	perc. diff BA	OE
OD_nDS_n	OD_2DS_2	DS+4	–	56	43	DS+13
	OD_3DS_3	DS+3	–	82	65	DS+17
	*OD_4DS_4	DS+2	–	83	45	DS+38
	*OD_5DS_5	DS+1	–	61	23	DS+38
OD_{n+1}- DS_n	OD_3DS_2	DS+3	DS -30	47	22	DS+25
	*OD_4DS_3	DS+2	DS -30	57	37	DS+20
	*OD_5DS_4	DS+1	DS -30	47	27	DS+20

The table is arranged in such a way that the "prominent" stimulus comes second in column one and two, and its respective values are given in column three, four and seven. First column: stimulus combinations. Second column: stimulus pairs in the order AB. Third column: difference in semitone between the members of the pair. The denoted stimulus differs from the other one by the indicated positive or negative value. Fourth column: difference in ms, analoguous to the third column. Fifth and sixth columns: percent different responses for the orders AB and BA. Seventh column: "amount" of OE as difference between columns five and six. Given is the value for the denoted ("prominent") stimulus.

Note that a thorough investigation of these aspects would have required a smaller spacing and much more experimental effort than we wanted to invest. These experiments should therefore be taken as pilot studies.

The stimulus combinations and the results are given in Table 2.

Some filler pairs occurred as well, namely "same" pairs and pairs which contained the edge stimuli 1 or 6. These pairs have been tested extensively in all other tests described till now and will not be discussed further. All stimulus pairs were repeated four times. 15 subjects participated in each experiment.

3.2.2.2 Results and discussion

The results for all pairs are given in columns 5 to 7 in Table 2. In the following discussion, we will only deal with the pairs marked with an asterisk, where we can be certain that clear rises were compared with each other, namely those pairs containing stimuli 4 and 5, cf. Figure 5. The motivation is that we are interested in the contribution of single acoustic factors to the OE and not in the ability of subjects to discriminate contour vs. level stimuli that are qualitatively different. In Figure 5, stimuli from the Offset-Slope Continuum are marked by a solid line, those from the Offset-Duration Continuum by a dashed line, and those from the Duration-Slope Continuum by a dotted line. The "amount" of the OE (seventh column in Table 2) is given for each combination as well; the prominent stimulus is marked by the letter "p".

The results can be summarized as follows:

(i) In pairs with the same Fo offset but different durations of the contour, the stimulus with the longer duration is more prominent (cf. pairs OD_4OS_4 and OD_5OS_5 in Figure 5). This result is comparable to those obtained for the isolated Duration-Slope Continuum (cf. part 3.2.1.).

(ii) In pairs having the same duration of the contour but different Fo offsets the stimulus with the higher Fo offset is more prominent (cf. OD_4DS_4 and OD_5DS_5).

(iii) Roughly, a difference of one semitone between the members of a pair can be compensated for by differences in the duration of the Fo contour of 30 to 60 ms and a more gentle Fo slope (cf. the pairs OD_4OS_3, OD_5OS_4, and OS_5DS_5 in Table 2 and Figure 5). However, the OE is more pronounced in pair OD_5DS_4, which has the same type of manipulation. This result might be caused by a steeper Fo slope (cf. Klatt 1973).

(iv) There is a tendency to compensate for a difference of 2 semitones with a duration, which is by 60 ms longer than in the stimulus having a higher offset (cf. pair OS_4DS_4).

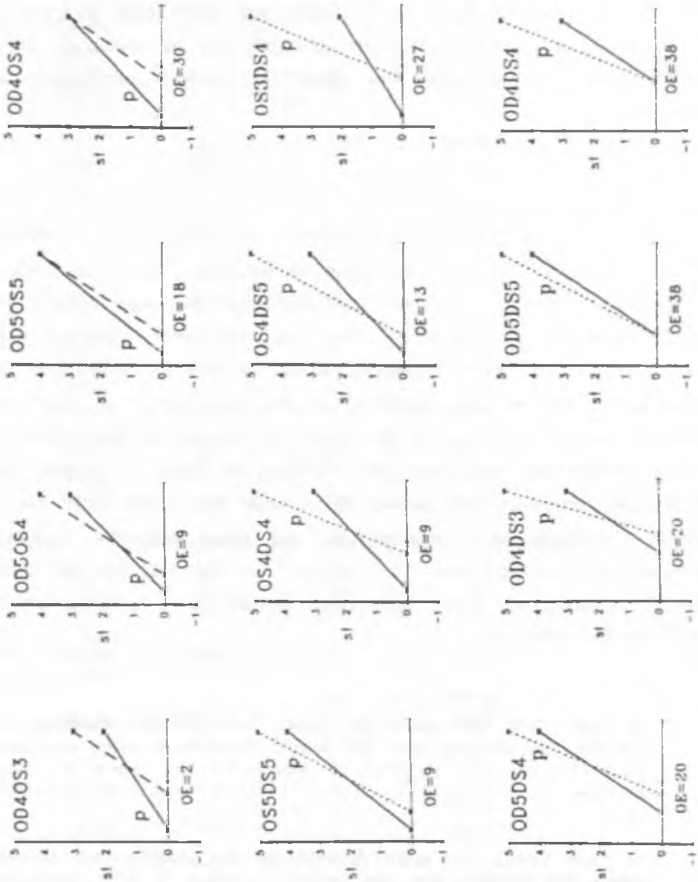

Figure 5: Mixed factors, single pairs

(v) A difference of three semitones cannot be compensated for by duration differences of 60 ms (cf. OS_3DS_4).

A multiple regression analysis was run in order to obtain a preliminary answer to the question of which factors are to which extent the relevant ones. As a measure of the goodness of fit, R^2 (the "explained variance") was chosen. For all 21 pairs, R^2 was rather low (.30); in a second step, only those 12 pairs marked with an asterisk in Table 2 (clear rises) were chosen. Here, the R^2 of .79 was sufficiently high; we can therefore speak of a strong dependence of the OE on the manipulated factors Fo offset and duration. The most relevant factor was the Fo offset with a univariate R^2 of .49, while for duration, the univariate R^2 was .26.

3.2.3 Discussion of the factors and preliminary conclusion

Our results concerning the relevance of the acoustic factors duration of the Fo contour, Fo offset, and Fo slope for the OE can be summarized as follows:

(i) If the members of a pair have the same duration of the contour, the stimulus with the higher Fo offset is the prominent one.

(ii) If the Fo offset of two stimuli is the same, the stimulus having the longer duration of the Fo contour is prominent, i.e., the stimulus with the more gentle slope.

(iii) If the Fo slope is identical, the stimulus with the longer duration of the Fo contour and higher Fo offset is more prominent.

(iv) Within a pair where the factors compensate each other up to a certain extent, the stimulus with a higher Fo offset is prominent, if the Fo offsets differ in two or more semitones.

3.3 DESIGN EFFECT

Thus far, the experiments have shown that the OE is no arbitrary phenomenon. It can be replicated, and conditions for an order to be prominent can be formulated. Nevertheless, we cannot be sure that it is not an experimental artifact that only occurs under certain experimental conditions as, e.g.:

(i) Presentation of both the prominent and the non-prominent order in the same test,

(ii) only one-step-size instead of, e.g., all-step-size,

(iii) manipulation of one single phonetic feature or a mixture of three features.

For an argumentation along similar lines, cf. Needham (1934a, 1934b, 1935).

In the following experiments, some conditions were varied which might have an influence on the OE.

3.3.1 Different randomization

3.3.1.1 Materials and procedure

With this group of experiments, we planned to investigate the possible influence of an anchoring effect from a different angle. Usually, the two orders, AB and BA, are mixed. If the OE is a surface phenomenon, which is substantially conditioned by the design, a separation of the two orders AB and BA into two separate blocks might produce results that are markedly different in comparison to the normal randomized condition.

For this task, a new Duration-Slope Continuum was generated with a smaller spacing of the interpolation and without a level stimulus; this is shown in Figure 6. With this continuum, we could ensure that there were more perceptually unequivocal rises in the continuum. The duration of the manipulated part was 40 ms for stimulus 1; it increased by 20 ms from stimulus 2 onwards. Because of the smaller spacing we chose a two-step design. Four tests were conducted:

(i) **Test 1 (BA-AB)**: The pairs with the non-prominent order BA occurred together with half of the "same" pairs in the first part of the test. In the second part, the pairs with the prominent order AB together with the other half of the "same" pairs followed. The first and the second parts were not separated by any pause, but the second followed without any interruption. There were five repetitions of each pair; 11 subjects took part.

(ii) **Test 2 (randomized)**: The items were randomized as usual, i.e., prominent and non-prominent orders were intermingled. There were five repetitions of each pair; 10 subjects took part.

(iii) **Test 3 (AB-BA)**: The order of the two blocks was the reverse of test 1, i.e., first came the prominent order AB, and then following in the second part, the non-prominent order BA. There were five repetitions of each pair; 15 subjects took part.

(iv) Test 4 (identification test): The difference of 5 semitones between onset and offset should cause the subjects to label the stimuli as questions instead of assertions, if, and only if, the duration of the rise is long enough so that they are able to perceive it at all. Besides, the category boundary can be compared with a peak of the function of the "same" and the "different" pairs. There were ten repetitions of each stimulus; 14 subjects took part.

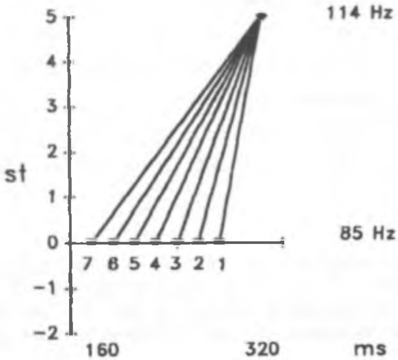

Figure 6: Duration-Slope Continuum, smaller spacing

3.3.1.2 Results and discussion

The results are given in Figure 7a for test 1, in Figure 7b for test 2, and in Figure 7c for test 3. The identification function for questions obtained in test 4 is given in each of the figures (dashed line). For test 1, the OE main effect ($F(1,14) = 75.77$, $p < .001$) and the stimulus pair effect ($F(4,14) = 9.81$, $p < .001$) are significant, the interaction is not ($F(4,14) = 2.22$, n.s.). The same holds for test 2: OE main effect: $F(1,10) = 79.83$, $p < .001$; stimulus pair effect: $F(4,10) = 6.99$, $p < .01$; interaction: $F(4,10) = 1.03$, n.s. For test 3, the OE main effect ($F(1,10) = 24.35$, $p < .001$) and the stimulus pair effect ($F(4.10) = 4.83$, $p < .05$) are significant, whereas the interaction is not ($F(4,10) = 0.60$, n.s.). In Figure 8, the AB and BA functions for all three randomizations are plotted together. A tendency can be observed that the prominent order AB is best discriminated in the BA-AB condition (test 1) and worst in the AB-BA condition (test 3). The non-prominent order BA is best discriminated in the BA-AB condition (test 1) and worst in the randomized condition (test 2). This means that the presence of the BA pairs (either in the randomized or in the BA-AB condition) increases the discrimination scores for the AB pairs, and vice versa, the presence of the AB pairs lowers the discrimination scores for the BA condition (mostly for the randomized condition).

These results show that the design can have a marked influence on the discrimination functions, but not on the OE. Three further points should be mentioned:

(i) The discrimination functions for AB and BA are similar in all three tests with a peak at pair 2/4, i.e., at stimulus 3. There is a tendency towards a plateau at stimuli 3 and 4. The plateau matches with the category boundary in the identification test between stimuli 3 and 4.

(ii) In the identification test, the subjects labeled stimuli 1 and 2 as assertions, and stimuli 4 to 7 as questions. Stimulus 3 is ambiguous. Obviously, the subjects need at least 100 ms rise time to perceive a stimulus as a clear question. This means that the peak of the discrimination functions, as well as the category boundary obtained with the identification test, are caused by psychoacoustic reasons: Discrimination between perceptually level and contour is better than between contour and contour.

(iii) The "same" pairs show a tendency towards a peak in the region of the category boundary between stimuli 3 to 5. This result can be interpreted along the same lines as in part 3.1.2.

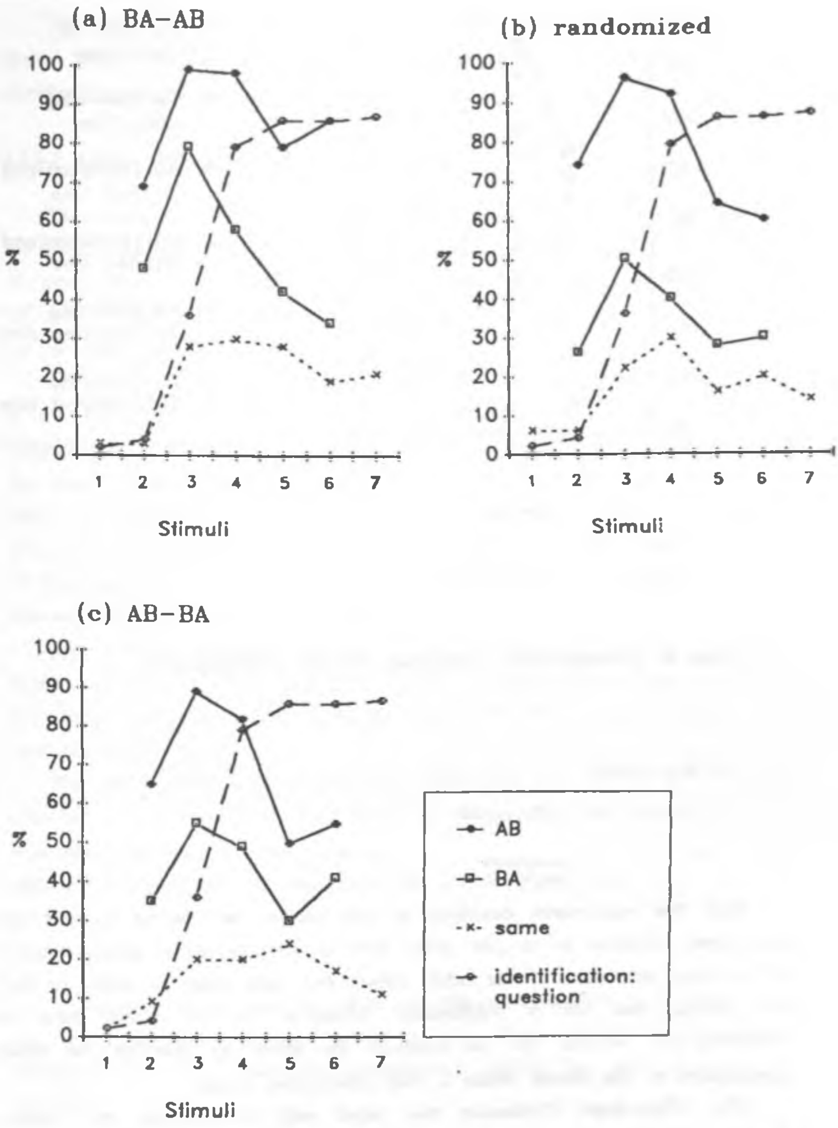

Figure 7: Duration-Slope Continuum, smaller spacing, discrimination functions

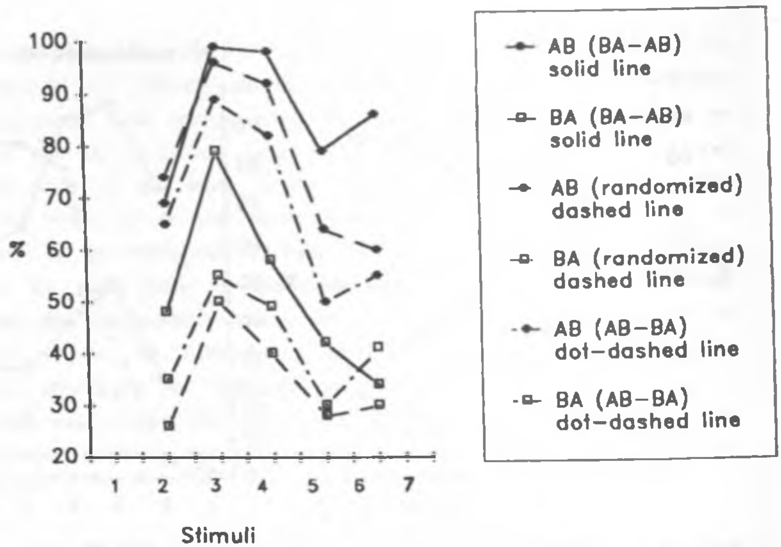

Figure 8: Duration-Slope Continuum, different randomizations

3.3.2 All-step design

3.3.2.1 Materials and procedure

With the experiments described in this section, we wanted to test the same three continua as in part 3.2.1. with an all-step design (every stimulus is combined with each of the other stimuli and with itself) in order to find out whether the OE is substantially influenced by this strong kind of anchoring, i.e., whether we can minimize the effect by changing the usual presentation of the stimuli within a fully randomized design.

The Offset-Slope Continuum was tested with 12 subjects, the Offset-Duration Continuum with 15, and the Duration-Slope Continuum with 14 subjects. Subjects were members of the staff and students of phonetics.

The default procedure was slightly modified. Since we wanted to avoid undue length of the experiments, there were only four repetitions of each combination instead of five.

3.3.2.2 Results and discussion

The results are given in Figure 9 and in Table 3. The four-step and five-step conditions produced a consistent ceiling effect with a discrimination of more than 97 %. They are therefore not taken into consideration. The OE main effect was significant for all three continua in the one-step and in the two-step condition.

In the Offset-Slope Continuum (Figure 9a), the same pairs have a peak in the region of stimulus 5, i.e., at the stimulus with a clear contour. A comparison of the one-step condition with the results of the experiment with the same stimuli, but with a one-step instead of an all-step condition, cf. Figure 4a, shows no marked difference, as far as the OE is concerned. However, due to an anchoring effect, the all-step design causes a lowering of the discrimination functions for the one-step pairs in comparison with the one-step design.

The resulting functions for the Offset-Duration Continuum are given in Figure 9b. A comparison with the results of the one-step test (cf. Figure 4b) shows no pronounced anchoring effect for the all-step condition (cf. especially pair 4/5).

For the Duration-Slope Continuum (Figure 9c), the OE for the three-step condition is not significant; very likely, a ceiling effect was already produced with that step-size. A slight anchoring effect can be observed in comparison with the one-step test (cf. especially pair 2/3 in Figure 4c).

3.3.3 General discussion of the design effect

A different design has an effect on the discrimination. The presence of the prominent pair AB (cf. 3.3.1.) or the clear anchors in the all-step condition (cf. 3.3.2.) lower the discrimination ability, but have no effect on the OE.

Table 3: Statistical results

test	step size	factor	df	F	p
Offset-Slope	1	OE	1,11	37.97	< .001
	1	Pairs	4,11	.86	n.s.
	1	Int.	4,11	4.28	< .05
	2	OE	1,11	51.56	< .001
	2	Pairs	3,11	2.18	n.s.
	2	Int.	3,11	3.09	n.s.
	3	OE	1,11	10.25	< .01
	3	Pairs	2,11	2.61	n.s.
	3	Int.	2,11	3.70	n.s.
Offset-Duration	1	OE	1,13	8.53	< .05
	1	Pairs	4,13	58.62	< .001
	1	Int.	4,13	1.91	n.s.
	2	OE	1,13	5.87	< .05
	2	Pairs	3,13	208.21	< .001
	2	Int.	3,13	2.29	n.s.
	3	OE	1,13	15.19	< .01
	3	Pairs	2,13	76.51	< .001
	3	Int.	2,13	10.30	< .01
Duration-Slope	1	OE	1,14	18.08	< .001
	1	Pairs	4,14	50.90	< .001
	1	Int.	4,14	1.16	n.s.
	2	OE	1,14	13.59	< .01
	2	Pairs	3,14	66.54	< .001
	2	Int.	3,14	3.41	< .05
	3	OE	1,14	3.12	n.s.
	3	Pairs	2,14	1.65	n.s.
	3	Int.	2,14	.78	n.s.

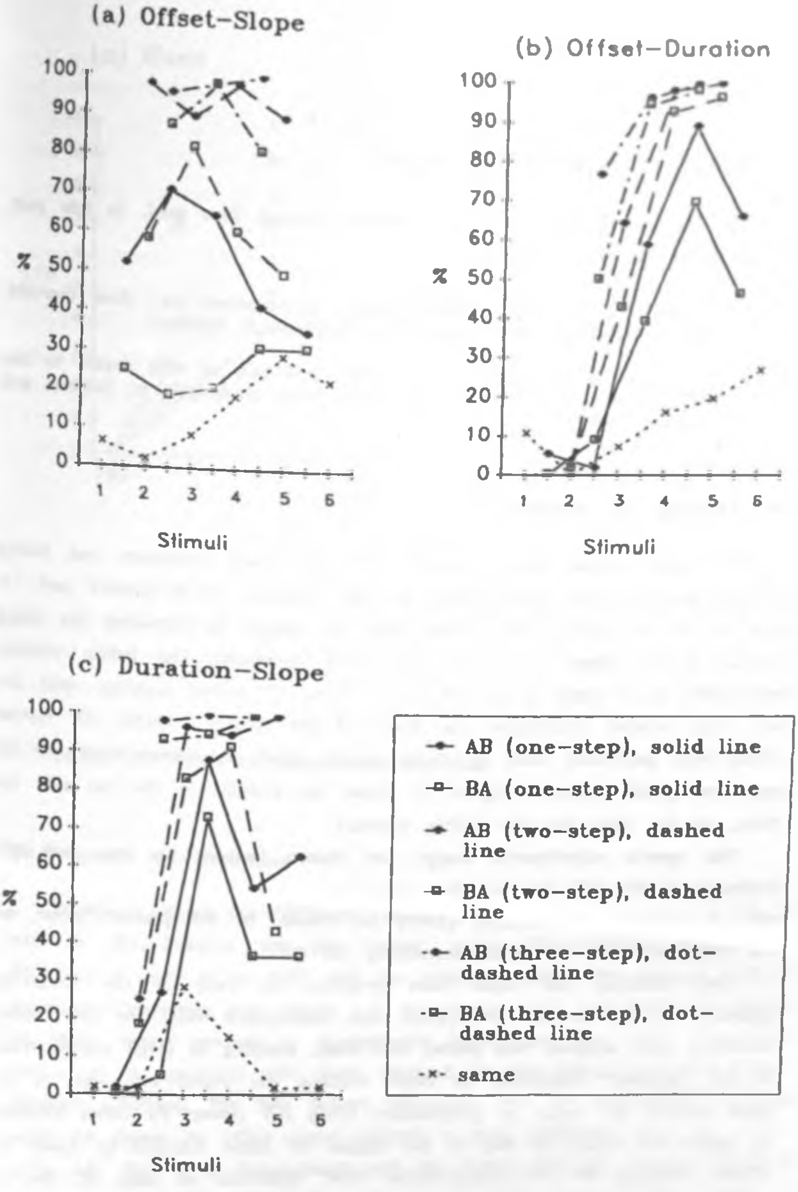

Figure 9: All step tests, discrimination functions

3.4 SPEECH VS. NON SPEECH MODE

Thus far, mostly rises and only speech material were used. In this part, we will focus on two further questions:

(i) Is the OE a purely psychoacoustic phenomenon, i.e., does the OE behave differently with speech and non-speech material?

(ii) Is there any difference between rises and falls as with regard to the OE? These questions were addressed more thoroughly in Batliner and Schiefer (1987).

3.4.1 Materials and procedure

Two target stimuli were generated with the usual procedure, one falling by one semitone, the other rising by one semitone in its second part. A total of 12 test stimuli were derived from the target by increasing the rising contour in six steps of a 1/8 tone and decreasing the falling contour analogously in 6 steps of a 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 to the respective speech stimulus. Figure 10 shows the stimuli, on the left side the rising, on the right side the falling contours.

This special experimental design was chosen because we were not only interested in the OE, but in other phenomena as, e.g., threshold detection as well, cf. Batliner and Schiefer (1987) for details. In the present paper, we will only deal with the results concerning the OE.

Four different test tapes were prepared for each of the subgroups (speech-rises, speech-falls, non-speech rises, non-speech falls). In the "same" condition, each stimulus was paired with itself, resulting in seven combinations. In the "different" condition, the target stimulus was paired with each of the other stimuli, the order of presentation being AB (lower Fo offset followed by higher Fo offset) as well as BA (higher Fo offset followed by lower Fo offset), resulting in 2*6 combinations. Five repetitions of each of the 19 combinations were taped in randomized order, with an inter stimulus interval of 500 ms between the members of a pair and with 3500 ms between the pairs. The number of subjects was 14 for the speech rises, 12 for the speech falls, 11 for the non-speech rises, and 14 for the non-speech falls.

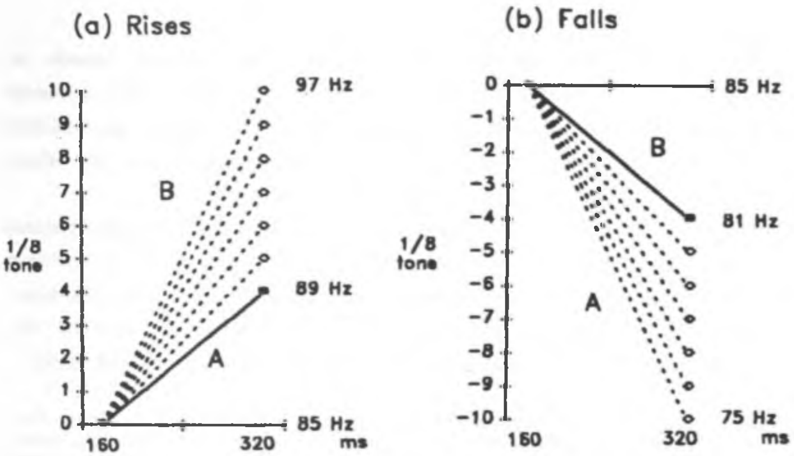

Figure 10: Target continua: Rises and Falls

3.4.2 Results and discussion

Figure 11 shows the "different" responses for the orders AB and BA, as well as the function for the "same" pairs. The abscissa displays the difference in tone (0 to 6/8), and the ordinate displays the percent different responses. Concerning the speech and non-speech rises, the order AB (higher Fo offset comes second) is discriminated better than the order BA. These results are in agreement with those obtained for the Offset-Slope and Offset-Duration Continua. The falls behave differently: For the non-speech falls the order BA (lower Fo offset comes second) is more prominent than the reverse order AB. This is in disagreement with the results obtained for the Rise-Fall Continuum (cf. Figure 2a). For the speech falls, none of the orders is prominent throughout: The functions AB and BA intersect each other.

A multivariate analysis of variance was applied to the different conditions of the four groups together with four factors, two of them being repeated measures, i.e., OE and stimulus pairs; the other two, material (speech vs. non-speech) and contour (rise vs. fall) were independent. The necessary assumptions for this approach were tested with the Cochran and Bartlett tests. Table 4 shows the F values and levels of significance for the effects tested.

Four of the tested effects turned out to be significant: Material, material by contour by OE, OE, and pair. As there was an interaction between material, contour, and OE, the significant main effect of OE cannot be interpreted. Therefore, Figure 12 displays the simple main effects for AB and BA. Two phenomena should be discussed which showed up clearly in this figure:

(i) Within the rises, the difference between AB and BA is more pronounced for speech than for non-speech material. A possible explanation might be that the combination of the more complex speech material with the non-prominent order BA causes more problems for the discrimination task.

(ii) Within the falls, non-speech material is discriminated better than speech material. As for the speech material, there is almost no difference between the orders AB and BA because of a crossing of the two functions (cf. Figure 11b). For the non-speech material, the order BA is discriminated much better than the order AB. The results for the speech falls are unsystematic. However, if we compare the non-speech falls (cf. Figure 11d) with the speech falls of the Rise-Fall Continuum (cf. Figure 3), then it looks as if the subjects followed two different strategies: In the case of non-speech material, the second, prominent stimulus shows a more pronounced Fo contour; in the case of speech material, the prominent stimulus shows a higher Fo offset.

We cannot really explain the results for the speech falls. It might be the special experimental design (target stimulus instead of fixed step-size as in the Rise-Fall Continuum). Alternatively, it might be that the subjects followed two different strategies, the one of them being rather a psychoacoustic, the other one rather a linguistic strategy, depending on the amount of the Fo change (cf. part 5). Obviously, the speech falls were more difficult to evaluate than the other stimuli, because the discrimination function is lower than in the other three experiments.

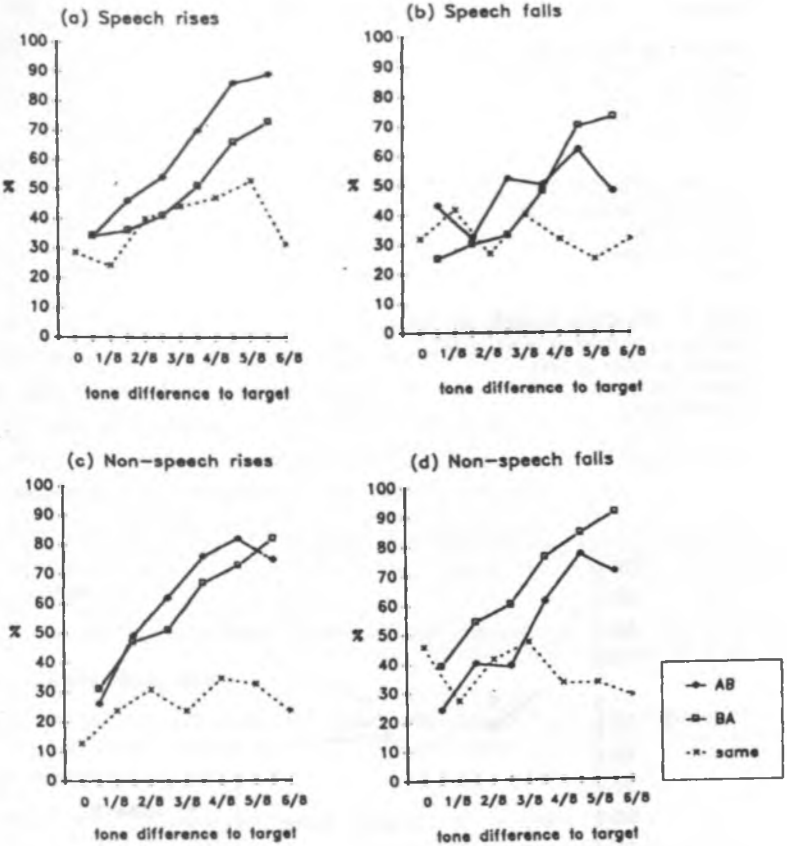

Figure 11: Spech and Non-Speech stimuli, discrimination functions

Table 4: Statistical results

	F	p <
BETWEEN-SUBJECTS (df: 1,47)		
material by contour	1.42	n.s.
contour	1.57	n.s.
material	4.22	.046
ORDER WITHIN SUBJECTS (df: 1,47)		
material by contour by order	6.95	.011
contour by order	.03	n.s.
material by order	.43	n.s.
order	9.14	.004
PAIR WITHIN SUBJECTS (df: 5,43)		
material by contour by pair	.87	n.s.
contour by pair	.60	n.s.
material by pair	1.96	n.s.
pair	29.09	.001
ORDER BY PAIR WITHIN SUBJECTS (df: 5,43)		
material by contour by order by pair	.35	n.s.
contour by order by pair	2.17	n.s.
material by order by pair	1.52	n.s.
order by pair	1.75	n.s.

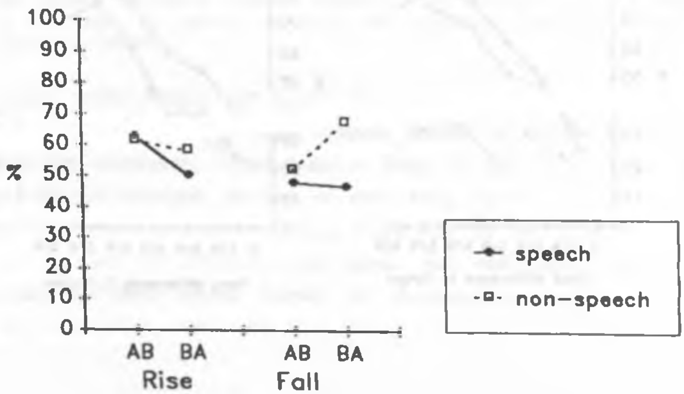

Figure 12: Interactions

3.5 SUMMARY OF RESULTS AND DISCUSSION

The results of the experiments lead to the following preliminary conclusions:

(i) The OE is no random effect, as it could be replicated in all the experiments.

(ii) The OE is not an experimental artifact that can be traced back to a special design. The design can influence the discrimination function, but it never influenced significantly the OE. (Of course, this conclusion holds only for the AX task used throughout our experiments.)

We now take it that the second question of part 1.3., whether the OE can be influenced by the experimental design, could be answered, but the other still remain: What are the factors that cause the OE? We would now like to look more closely at the phonetic parameters of our stimuli. If we take the second stimulus in the prominent order (i.e. the prominent stimulus) and compare it with the first, then we can say that

(i) in the Offset-Slope Continuum, this stimulus has same duration of the Fo contour, but higher offset, greater Fo mean and steeper slope,

(ii) in the Duration-Slope Continuum, this stimulus has same offset and same Fo mean, but longer duration of the Fo contour and a more gentle slope, and

(iii) in the Offset-Duration Continuum, this stimulus has the same slope, but longer duration of the Fo contour, higher offset and greater Fo mean.

The OE is therefore not simply triggered by a single (acoustic) factor. Yet, there could exist a single, non-acoustic factor that might explain these results.

4. PHONETIC AND PSYCHOACOUSTIC EVIDENCE

In part 4.1., we shall report on phonetic work which might help to explain the prominence of a stimulus that is caused by either a higher Fo offset or a longer duration of the Fo contour. The prominent order is addressed in part 4.2. Based on these findings, a preliminary explanation of our results is given in part 4.3.

The boundaries between two fields as, e.g., psychoacoustics and phonetics or between the different domains of phonetics itself are difficult to define. A "simple" convention might be sufficient for this paper: We will distinguish a "psychoacoustic/phonetic" domain and a "phonetic/linguistic" domain. The former comprises the perception of non-speech sounds (psychoacoustics) and of speech sounds (phonetics) as well as their production and articulation; the latter comprises the linguistic, i.e., the conventionalized use of speech sounds.

4.1 THE "PROMINENT" STIMULUS

4.1.1 Evidence from tone languages

In tone languages, it has been found that

(i) syllables with high tones are shorter compared with those having a mid or low tone (cf. Cantonese (Benedict 1948), Zapotec (Pike 1948), Chatino (Pike 1974), Athapaskian (Blight and Pike 1976), and Thai (Abramson 1962)), and

(ii) syllables with falling tones are shorter than those with rising tones (cf. Yuman (Langdon 1976), Thai (Abramson 1962), and Mandarin (Dreher and Lee 1966)). Gandour (1977:60) summarizes the results from tone languages as follows: "Other factors being equal, (a) vowels (syllables) on low tones are longer than those on high tones; (b) vowels (syllables) on rising tones are longer than those on falling tones, and (c) vowel (syllable) duration is inversely related to the approximate average fundamental frequency".

That these results are far from random but, on the contrary, reflect universal phonetic tendencies in the production of tones can be seen if one takes diachronical processes into consideration: E.g., in Thai dialects, a historical loss of the phonological distinction in vowel length can be observed. In one dialect, Chiang Rai, "... short non-low vowels have become long under rising tones, long non-low vowels have become short under nonrising tones" (Gandour 1977:57). In another dialect, Phuket, "... long non-low vowels have become short under falling tones, all short vowels have become long under nonfalling tones" (Gandour 1977:58). If we do not take into consideration the tongue height and the difference in clustering of the vowels, we note the same effect of tones on vowels in both dialects: Falling tones lead to a shortening and rising tones to a prolongation of vowels.

Additional evidence could be vowel lengthening under circumflex (rising) accent in Lithuanian, (cf. Kenstowicz 1972:4ff and Halle and Vergnaud 1987:200).

4.1.2 Evidence from production experiments

In order to investigate a possible interrelationship between the Fo contour of a tone and its duration, some production experiments were undertaken in which subjects had to produce tones within a special experimental design. Leonard and Ringel (1979), for example, conducted an experiment on vocal shadowing using square-wave modulated complex tones. Their results show that pitch lowering is faster than pitch rising.

Sundberg (1979) examined the maximum speed of pitch changes in singers and untrained subjects. Two given tones had to be reproduced rhythmically as fast as possible. The frequency ratio between both tones was 1:2, 1:1.5, 1:1.25, 1:1.12. The results showed that "The response time is shorter in pitch drops than in pitch elevations, in small pitch changes, in female subjects, and in trained subjects." (Sundberg 1979:74). There is an interdependence between the direction of the pitch change and the response time in untrained subjects, as pitch elevations take longer than pitch drops.

The differences in the adjustment time between small and large tonal intervals can be explained by the tension of the laryngeal muscles and the subglottal air pressure needed for the realization of a given tone. A greater tension of the muscles and a higher subglottal air pressure is required for higher tones. The greater the difference between two tones, the greater the required effort, which results in a greater time interval needed to achieve a given tone.

A mere relaxing of the laryngeal muscles, i.e., a purely passive mechanism, can presumably not explain the remarkable discrepancies between the speed of pitch rises and pitch falls in Leonard and Ringel (1979) and Sundberg (1979; cf. especially the results for the untrained subjects). In order to find an explanation of the results, Leonard and Ringel as well as Sundberg refer to Ohala (1972) and Ohala and Ewan (1973), who argue that pitch is lowered by active mechanisms supplementing the passive ones.

Sundberg (1979) claims that two different physiological mechanisms are involved, one in frequency rising and the other one in frequency lowering. He assumes that the main muscle in frequency lowering is the thyro-arytenoideus lateralis. This muscle is also known as one of the adductor muscles and thus serves as a protector of the larynx and the lungs (Hardcastle 1976). It is well known that protector muscles are well developed and fast responding because they have a vital function. Pitch rising, on the other hand, is mainly achieved by the crico-thyroid muscle, which is believed to have no protecting function (Hardcastle 1976:81). However, this muscle may behave differently depending upon the subject's training. The differences between the trained and untrained subjects in pitch elevations vs. pitch drops in Sundberg (1979) can thus be explained. There is no significant difference between the groups of subjects for pitch drops, as the thyro-arythenoideus lateralis is, even without any training, fast responding in both groups of subjects, whereas rises are performed more quickly by singers due to the training of the crico-thyroid muscle.

The results from the production experiments correspond very well with those from tone languages, where it was found that rising tones are longer than falling tones. It follows from Leonard and Ringel (1979) and Sundberg (1979) that greater pitch intervals require a longer adjustment time and that falls are performed more quickly than rises. If we take these results as phonetic universals, we can assume that greater pitch intervals are always connected articulatorily and auditorily with greater durations, and vice versa, greater durations of pitch elevations or pitch drops are related to a greater amount of pitch change and - in the case of rises - a higher or - in the case of falls - a lower Fo offset.

4.1.3 Evidence from perception experiments

The difference in the behavior of subjects towards rises and falls has been replicated by several perception experiments in which speech stimuli were used carrying certain Fo contours as, e.g., level, rising, and falling contours (cf. Klatt 1973, 't Hart 1981, and Black 1970). From these experiments we know that level tones are better discriminated than contour tones. Klatt (1973), for example, could show that whereas only a difference of 0.3 Hz between two level tones is required to yield a "different" response, stimuli with falling contours (and the same slope) required 2.0 to 2.5 Hz. Contour tones with a steeper slope (32 Hz/sec) needed a greater difference (4.0 Hz) than those with a less steep slope (12 Hz/sec). Contour tones of the same direction are easier to discriminate than those of different directions.

't Hart (1981), who used speech stimuli as well as piano tones found remarkable differences between subjects, some of which were not able to perceive any difference if the tones were separated by less than 4 semitones. Those who performed well on the task were better able to judge speech rises than speech falls.

Black (1970:180) reports the results of different perception experiments: "Irrespective of the psychophysical method, slow inflections were of greater impact than fast ones of the same extent, and upward inflections were generally more powerful than downward ones of the same extent and rate. ... Long inflections were generally judged to be of greater magnitude than short ones of similar rate. The effect of short-slow inflections was quite comparable to that of long-fast ones."

These results reported in the literature suggest that the perceptual effect of a higher Fo offset might be equal to that of a longer duration of a Fo contour, as both factors are normally interrelated. In our experiments, however, a longer lasting elevation of Fo (longer duration) does not lead to a higher Fo offset as both factors were handled independently. At any rate, subjects seem to perceive a higher Fo offset, if the Fo contour is longer and, vice versa, a lower Fo offset if the Fo contour is shorter. This is indicated by the results from the "mixed" experiments (part 3.2.2.), where Fo offset and duration of the Fo contour are in trading relation (cf. Lehiste 1976). If we assume that articulatory gestures are "perceived" by the listener, then the prominence of a stimulus is caused by a greater effort in the production, i.e., a higher muscular tension needed to achieve a steeper rising or falling Fo contour and a higher or lower Fo offset as well. This assumption does not hold true for the speech falls reported on in part 3.1.2., where stimuli with a higher Fo offset, not with a greater amount of Fo change, are likely to be perceived as prominent. The explanation of this phenomenon has to be postponed till part 5.

The prominence of a stimulus can thus be explained by articulatory and/or physiological mechanisms. The prominence of a stimulus order will be dealt with in the next part.

4.2 PROMINENT ORDER

The crucial point in the evaluation of two events A and B that follow each other on the time axis was already mentioned in part 1.2: Item A has to be kept in the memory and then compared with item B. In our material, the manipulated variable was Fo, the evaluation time is surely not below 320 ms (end of stimulus A) and not much above 1140 ms (end of stimulus B). The prominent stimulus, as established in part 4.1 always comes second in the prominent order; note again that we will postpone an explanation for the speech falls till part 5. It is evident that the information concerning stimulus A was not lost for our subjects at evaluation time - otherwise, no systematic results could have been achieved - it is evident as well that something must have happened to stimulus A - otherwise, there would be no order AB that is more prominent than the order BA. Some sort of phonetic/linguistic encoding (labeling/categorization) might have taken place at evaluation time, cf. especially the peak of the "different" and the "same" functions at the category boundary; still, item A must have been represented in an image-like mode as well, cf. especially the roughly linear functions in Figure 11.

Most models of memory organization assume a two-stage processing of acoustic information. First, some sort of acoustic buffer (auditory memory) where the signal is stored in a relatively unaltered form; second - especially in the case of speech signals - a (short term) memory, where the signal is already decoded into (speech specific) information (phonemes, words, etc.). There is no agreement as for the time domain of the auditory memory; obviously, it depends on the specific information to be stored: Vowels, e.g., seem to have a longer decay time than consonants. It is claimed that auditory memory for pitch may persist relatively unaltered over intervals of more than 2 sec up to 15 sec; cf. for the state of the art Helfrich (1985:280ff), and for results concerning pitch especially Wolf (1977) and Deutsch (1975). But " ... memory for tonal pitch is subject to a large interference effect caused specifically by other tones and not due to some general storage limitations." (Deutsch 1975:113).

Let us now try to bring our results in relation to the state of the art. We can conclude that at evaluation time the Fo information of stimulus A is still kept in memory, but that it is influenced by the Fo information of stimulus B. If we substitute "weakened" for "influenced", then the "prominent" order can be explained: The auditory trace of stimulus A is weakened by stimulus B.

An explanation along similar lines was already offered by Köhler (1923), cf. his "Absinken der Spur" (Köhler 1923:158), i.e., "fading/decay of the (auditory) trace". Chuang and Wang (1978) report on comparable results; in their experiments, there is a tendency of "overestimating the pitch of the second sound in temporal pitch comparison" (Chuang and Wang 1978:1004) if two stimuli with the same Fo course have to be compared. This means, in physically identical stimulus pairs, the second stimulus will be perceived as being higher in pitch. The same holds true for duration, if it exceeds a given threshold. These results match very well with our own findings, as stimulus pairs containing identical stimuli are often judged as different. They also agree with the results from the different pairs: If the pitch of the second stimulus is overestimated, the perceptual difference between the first and second stimulus increases, if the second one is higher in pitch and/or longer in duration. In the case of BA pairs, the perceptual difference decreases.

4.3 A PRELIMINARY EXPLANATION AND ITS LIMITATIONS

In Figure 13, the following stimulus pairs are represented schematically in the combinations AB and BA (B always being the prominent stimulus): Rises from the Rise-Fall Continuum (identical with the stimuli from the Offset-Slope Continuum), rises from the Duration-Slope Continuum, and falls from the Rise-Fall Continuum. The "weakening" of the first stimulus is indicated by an asterisk; in the case of the rises, it leads perceptually to a greater difference in the combination AB and to a smaller difference in the combination BA. As for the rises, the OE in our experiments can thus be explained. The results for the non-speech falls, cf. Figure 11d, can be explained as well by this sort of "weakening", but not those for the speech falls in the Rise-Fall Continuum, cf. Figure 3. There, the weakening of the first stimulus should result in a less pronounced difference in the case of AB (prominent order in the Rise-Fall Continuum), and vice versa.

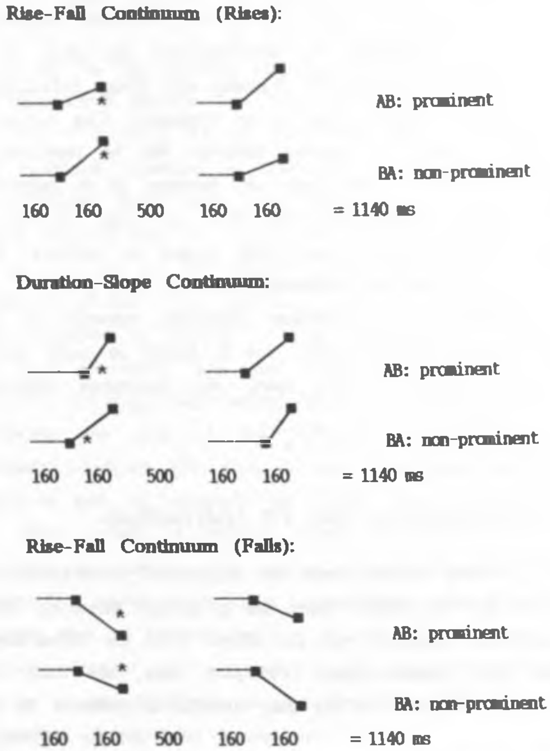

Figure 13: Prominent vs. non-prominent order

5. ORDER EFFECT AND THE ORDER OF ACCENTS

As no phonetic/psychoacoustic explanation for the behavior of the speech falls in the Rise-Fall Continuum can be given, it might be worth while to look at the differences in the processing mode (speech vs. non-speech mode of perception). In this part 5, we will therefore search for an explanation of the OE in the phonetic/linguistic domain.

5.1 ORDER EFFECT AND REAL LIFE

Up till now, we took the OE for an experimental phenomenon, and we did not discuss the possibility that it has relevance for "real life". One could really consider the task of comparing two events one after the other and classify them as "same" or "different" as a rather academic endeavour - and if not academic, at least seldom and conscious. Normally, we do not often compare two weights, such a situation might come forth, however, when comparing two mountainbikes. But, if we are not a salesperson, this does not happen every day. Nor do we very often "compare" two speech events in the literal meaning; this is the task of linguists or phoneticians, but not of "normal" people. Nevertheless, there is at least one task for the "normal" native speaker/hearer that is comparable to the task of our subjects and that he/she has to accomplish in everyday conversation: To decide which of the phrases in an utterance carries the focal accent (FA) and thereby the "new" information. Quite often this decision is based on syntactic/semantic/pragmatic factors or, if in doubt, on a default position of the FA; but it can be based only on the acoustic features of the phrases in question as well.

In the following sections, we will therefore summarize the results of an investigation (cf. Batliner 1989b, Batliner and Nöth 1989) on the phonetic structure of the FAs in German, and then compare these results with the OE.

5.2 PROMINENCE OF ACCENTS IN GERMAN

The material consisted of 360 German utterances, spoken by six untrained speakers (3 male, 3 female). Three different sentences with a similar syntactic structure were each put in different contexts that determined sentence modality as well as place and manner of focus; for a detailed description of the corpus and the intended focal structures cf. Batliner and Oppenrieder (1989) and Oppenrieder (1989). In each of the sentences, the last two phrases could be stressed depending on the surrounding context. The sentences formed minimal pairs that could only be differentiated by their intonational form: FA in final vs. FA in prefinal position, on the one hand, and questions (Qs) vs. non-questions (NQs), on the other hand. Table 5 shows the three test sentences, an English translation, and a finer description of the induced sentence modalities Q/NQ, NQ being either assertion, imperative, or adhortative.

Table 4: Test sentences, translation, and induced sentence modalities

Sie läßt die Nina das Leinen weben ?/.	
She makes Nina weave the linen	
assertive question vs. assertion	
Lassen Sie den Manni die Bohnen schneiden ?/!	
Make Manni cut the beans	
polar question vs. imperative	
Lassen wir den Leo die Blumen düngen ?/!	
let us make Leo fertilize the flowers	
polar question vs. adhortative	

The only instruction given to the speakers was to produce the context and the test sentence. We did not instruct the speakers to produce the FA or Qs/NQs in a certain way. The position of the FA that our analysis is based on was not determined simply by the intended position that, in turn, was indicated by the context and, hopefully, realized by the speakers. Instead, perception experiments were run where the sentences were presented in isolation to groups of listeners (12 subjects on the average). These listeners had to decide on the position of the FA; for details, cf. Batliner and Nöth (1989:211). The task of these subjects to decide upon the most prominent phrase is thus comparable to the task in a "same-different"-task: No contextual information whatsoever is given.

Each of the 4 Q/NQ-FA constellations has one central type that is characterized by the average values of the relevant intonational predictor variables. (For details, cf. Batliner 1989b and Batliner and Nöth 1989).

For these 4 central types, Figure 14 shows the average feature values as well as the Fo contour of a prototypical production; prototypicality was assumed if the utterances passed several criteria, e.g., could be classified as very natural productions, etc.; cf. Batliner 1989b:56ff.

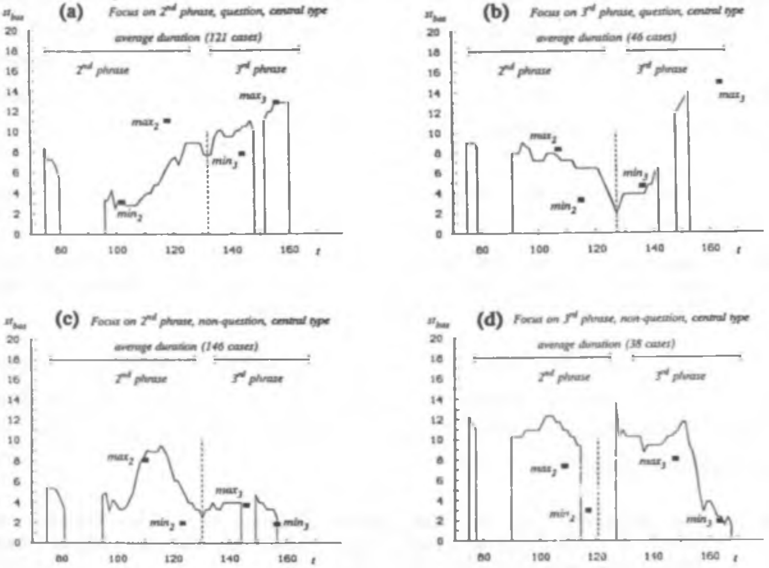

Figure 14: Focal Accent material

The dashed vertical line in Figure 14 marks the border between the 2nd and the 3rd phrase of the actual production. For the 2nd and the 3rd phrase (first or second position of the FA), each of the filled squares shows averages for max_2, min_2, max_3 *and* min_3, i.e. Fo maxima and Fo minima on the 2nd and the 3rd phrase respectively. The position on the abscissa corresponds to the average position on the time axis in centiseconds starting from the beginning of the utterance; the position on the ordinate corresponds to the average Fo values in semitones above the speaker-specific basic value (st_{bas}). On top of each figure, average beginning point and duration of the 2nd and 3rd phrase is displayed; the number of cases is given as well.

In the following characterization, the terms "High", "Low", and "boundary tone" (cf. the tone sequence model, e.g., in Pierrehumbert 1980) are used interchangeably with the terms "rising"/"falling" contour:

(i) **Questions, FA on 2nd phrase** (Figure 14a): The contour is rising in both phrases (Low High).

(ii) **Questions, FA on 3rd phrase** (Figure 14b): In the 2nd phrase, this type has a falling contour comparable to the NQs, whereas in the 3rd phrase, the contour is rising (Low High).

Comparing these two types, we can say that the Fo range of the phrase with the FA is markedly greater than that of the other phrase. In the final phrase, a rising contour (high boundary tone) is used for both types to mark sentence modality.

(iii) **Non-Questions, FA on 2nd phrase** (Figure 14c): The contour is falling in both phrases (High Low). Max_2 is markedly higher than max_3; min_2 and min_3 do not differ.

(iv) **Non-Questions, FA on 3rd phrase** (Figure 14d): The contour is again falling in both phrases (High Low). Max_3 is about as high as max_2, min_2 and min_3 do not differ.

Comparing the two types, we can say that the absolute values for the features of the 2nd phrase in Figures 14c and 14d do not differ remarkably. It is rather the relative values of the features in comparison with the respective values of the 3rd phrase that mark the Fa.

For Qs, the most relevant factors are the Fo maximum on the second phrase and the position of this maximum on the time axis; for NQs, Fo maximum and duration of the third phrase are most relevant.

5.3 ORDER EFFECT, FOCAL ACCENTS, AND SPEECH VS. NON-SPEECH MODE OF PERCEPTION

In Figure 15, a sort of overlay plot of Figure 13 with Figures 14a to 14d is shown, i.e., the schematic description of the prominent vs. the non-prominent order (rises and falls in the Rise-Fall Continuum) is compared with the average values of the FAs. The scales of the abscissa and the ordinate are analogous to those in Figures 14a to 14d: Semitones and centiseconds. The OE stimuli (open circles) are aligned to the FA stimuli (full squares) in the following way:

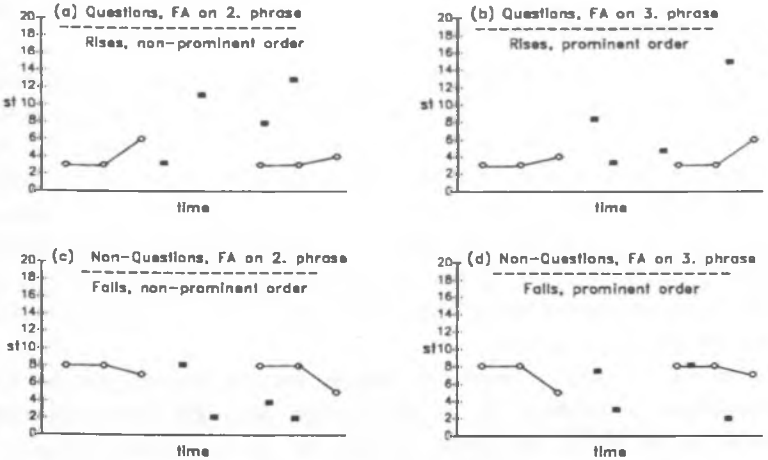

Figure 15: Overlay plot of FA material and OE material

(i) The height of the unmanipulated level part is aligned to the height of the turning points in the FA material (Fo minima in Qs, Fo maxima in NQs).

(ii) On the time axis, the beginning of the stimuli in the OE material is aligned to the beginning of the second phrase in the FA material. It can be seen that the duration of one *ja* differs not all that much from the duration of one phrase, but, of course, the 500 ms inter stimulus interval is missing in the FA material.

(iii) In order to show the tendencies more clearly, the Fo offsets of the OE stimuli differ by two semitones in each pair. (Note that we did not run an experiment for the Rise-Fall Continuum with such a two-step task.)

In some aspects, the OE material and the FA material cannot be compared in the strict sense. The "turning point" in the OE material was fixed on 84 Hz, whereas in the FA material, the turning points, i.e., the maximal Fo values in NQs and the minimal Fo values in Qs can be varied to a great extent, because they are the most relevant features for the marking of the FA.

As for the Q/FA constellation and the OE rises in Figure 15, the point of comparison is the more pronounced rise on the prominent stimulus in the OE material and on the phrase that carries the FA; this phrase can be called prominent as well. The prominent order AB, where the prominent stimulus comes second, corresponds to a FA on the third (last) phrase. (Note that for Q/FA on 3rd phrase, a fall is displayed on the second phrase, because only the greatest Fo movement per phrase is shown; it follows that the amount of the rise on the second phrase must be even less than the amount of the fall.)

As for the falls, a discrepancy between the OE material and the FA material can be observed. In the FA material, again the more pronounced fall is on the phrase that carries the FA, but in the prominent order AB of the Rise-Fall Continuum, the prominent stimulus B has a less pronounced fall than the non-prominent stimulus A. We believe that a solution can be found if we take the two stimuli that follow each other *(ja-ja)* not only as two acoustic or "purely" phonetic (i.e. auditory/articulatory) events but as some linguistic "gestalt" analogous to an utterance produced by a "normal" native speaker and perceived by a "normal" native listener. The most relevant variable for a classification of place of FA in NQs is the Fo maximum on the third phrase, cf. Batliner (1989b:39ff). This "Fo maximum" (i.e. the turning point) was held constant in the OE material. The stimuli in the OE material consist of one syllable each *(ja)*, whereas in the FA material, there is one syllable that can carry the accent followed by an unaccented syllable (e.g. *LEInen* and *WEben*, resp.); i.e. the speaker can produce a "fully developed contour" on the phrase that carries the FA. In Figure 16, one actual contour is displayed for each of the four Q/NQ-FA constellations for the utterance *Der Leo säuft* (as for more details, cf. Batliner and Oppenrieder 1989:322). For NQs, the Fo maximum on *säuft* is as high as that on *Leo*, if *säuft* is in focus. It is plausible that the most important feature (Fo maximum on the 3rd phrase) must be pronounced, whereas the amount of the fall is subject to variation dependent on the syllable structure. The Fo offset covaries with the height of the Fo maximum if only one syllable is the domain, but not in the case of a two-syllable utterance. It is surely possible to produce a fully developed contour even on *ja* and on *säuft;* this contour, however, is marked and not the default case. (Note that this conclusion should of course be checked against other material.)

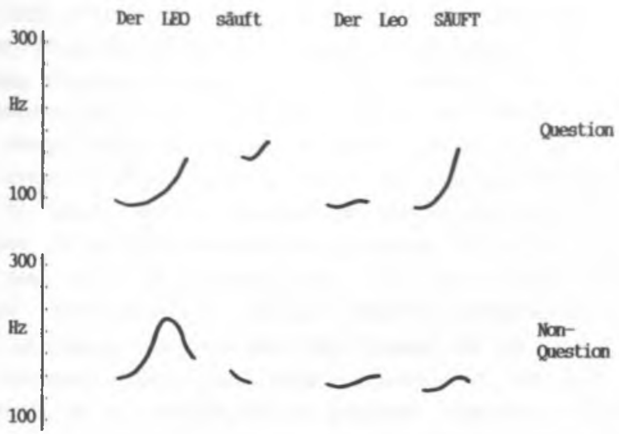

Figure 16: Focal accent, shorter time domain

The declination line covaries as well with the height of the Fo maximum, irrespective of the special computation (all point regression line, or a regression line through the Fo minima). If there is "enough space" (i.e., if the domain of the contour is two or more syllables), the declination line is falling, if not, it is rising. If we imagine a (speech specific) declination line (for the sake of the argument, an all point regression line) then, in the case of the FA on the 2nd phrase and in the case of the order BA, the declination line is steeper than in the case of the FA on the 3rd phrase and in the case of the order AB. Ceteris paribus, a rather flat (or rising) declination line indicates openess and/or prominence on the final part of the utterance. The effect is the same: The second item is more prominent, even if the amount of the Fo movement is less! (Note that we do not necessarily plead in favor of a declination line as the decisive "underlying entity"; it merely seems to be the most convenient way to sum up the traits in common.)

6. FINAL DISCUSSION

We have found that one order can be better discriminated than the other one; this was called the "prominent order". Phonetic/psychoacoustic reasoning lead us to the conclusion that in the prominent order, the second stimulus is more prominent than the first one, c.f. part 4.1. The concept of "prominence" is the link to the marking of the FA in natural speech. In the material described in part 5.2., there are two possible locations of prominence, as is the case in the OE material. The Fo contour of the prominent stimulus in the OE material can be compared with the Fo contour of the FA of the third phrase in the natural material. As for the rises, the interpretation is straightforward. Phonetic, linguistic, and psychoacoustic factors cannot be told apart. For the (speech) falls, some additional assumptions have to be made that can be summarized under the heading "perception of linguistic gestalt". A reasonable conclusion can be expressed in the following: Generally, the OE is a phonetic/psychoacoustic phenomenon, but under some circumstances, the (conventionalized) speech mode of perception might override the impact of the phonetic/psychoacoustic factors.

The caveat has to be made that in our experiments we concentrated on the rises, actually, for the simple reason to keep the effort in reasonable limits. It looks as if this concentration did pay off: The explanation offered might stand up to further examination. The falls were not tested extensively; yet, or maybe because of that very fact, they turned out to be an even more interesting object of investigation. Of course, we hope that our explanation, which heavily depends on the difference between speech and non-speech mode of perception, is valid as well. It goes without saying that at least the same experimental effort, as for the rises, should be spent on the falls in order to confirm this hypothesis. Irrespective of the validity of this special hypothesis, in the long run, the investigation of the OE might, vice versa, contribute to a better understanding of accent phenomena.

7. EPILOGUE

Actually, we did not finish with verified hypotheses, but rather with a program for future work. Progress is relative; though we hope that we do know a bit more about the OE than Fechner did 130 years ago, the problem is not yet solved. We started with a quotation by Fechner, and he might as well have the final say: " ... und noch heute, nachdem ich lange im Gebiete, namentlich des Gewichts- und Tastmasses, über dieselben [= die konstanten Fehler, i.e., the OE] experimentirt habe, ist mir der letzte Grund derselben grösstentheils unklar, und nur die Thatsache derselben sicher." (Fechner 1860 [1964]: 91). [... up till now, I have often investigated the OE. Still, I am looking for an explanation. It is only the very existence of this effect I am sure about.]

REFERENCES

Abramson, A.S. (1962): The vowels and tones of Standard Thai: acoustical measurements and experiments. *Bloomington, Indiana University, Research Center in Anthropology, Folklore, and Linguistics* 20.

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

Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): *Zur Intonation von Modus und Fokus im Deutschen.* Tübingen 1989.

Batliner, A. (1989a): Wieviel Halbtöne braucht die Frage? Merkmale, Dimensionen, Kategorien. In: Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): 111-162.

Batliner, A. (1989b): Fokus, Modus und die große Zahl. Zur intonatorischen Indizierung des Fokus im Deutschen. In: Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): 21-70.

Batliner, A. (1989c): Eine Frage ist eine Frage ist keine Frage. Perzeptionsexperimente zum Fragemodus im Deutschen. In: Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): 87-109.

Batliner, A. and Nöth, E. (1989): The prediction of focus. In: Tubach, J.P. and Mariani, J.J. (eds.): *Eurospeech 89.* Paris: 210-213.

Batliner, A. and Oppenrieder, W. (1989): Korpora und Auswertung. In: Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): 281-330.

Batliner, A. - Schiefer, L. (1987): Stimulus category, reaction time, and order effect - An experiment on pitch discrimination. In: *Proc. 11th ICPhS, Vol. 5.*: 46-49.

Benedict, P.K. (1948): Tonal systems in Southeast Asia. *J. Am.Orient.Soc.* 68: 184-191.

Black, J.W. (1970): The magnitude of pitch inflection. *Proc.6th ICPhS,* München: 177-181.

Blight, R.C. and Pike, E.V. (1976): The phonology of Tenango Otomi. *Int. J. Am. Ling.* 42: 51-57.

Chuang, C.-K. and Wang, W.S.-Y. (1978): Psychophysical pitch biases related to vowel quality, intensity difference, and sequential order. *J. Acoust. Soc. Am.* **64**: 1004-1014.

Deutsch, D. (1975): The organization of short-term memory for a single acoustic attribute. In: Deutsch, D. and Deutsch, J.A. (eds.): *Short-term memory.* New York, etc.: 107-151.

Dreher, J.J. and Lee, P.C. (1966): Instrumental investigation of single and paired Mandarin tonemes. *DARL Res. Commun. No. 13* (Adv.Res.Lab, Douglas Aircraft Company, Humlington Beach).

Fechner, G. Th. (1860): *Elemente der Psychophysik.* Leipzig. [1964 reprinted. Amsterdam].

Fowler, C.A. (1986): An event approach to the study of speech from a direct-realist perspective. *J. Phonetics* **14**: 3-28.

Gandour, J. (1977): On the interaction between tone and vowel length: evidence from Thai dialects. *Phonetica* **34**: 54-65.

Guttman, L. (1977): What is not what in statistics. *The Statistician* **26**: 81-107.

Halle, M. and Vergnaud, J.-R. (1987): *An essay on stress.* Cambridge, Mass. – London.

Hardcastle, W.J. (1976): *Physiology of speech perception.* London, etc.

Helfrich, H. (1985): *Satzmelodie und Sprachwahrnehmung.* Berlin – New York.

Hombert, J.-M, Ohala, J.J. and Ewan, W.G. (1979): Phonetic explanations for the development of tones. *Language* **55**: 37-58.

Jamieson, D.G. and Petrusic, W.M. (1975): Presentation order effects in duration discrimination. *Perception & Psychophysics* **17**: 197-202.

Kenstowicz, M. (1972): Lithuanian phonology. *Studies in the Linguistic Sciences 10.2.* (Department of Linguistics, University of Illinois, Urbana).

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

Köhler, W. (1923): Zur Theorie des Sukzessivvergleichs und der Zeitfehler. *Psycholog. Forschungen* 4: 115-175.

Langdon, M. (1976): Stress, length and pitch in Yuman languages. *Proc. Stressfest* (University of Texas Press, Austin).

Lauenstein, O. (1932): Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler. *Psycholog. Forschungen* 13: 130-177.

Lehiste, I. (1976): Influence of fundamental frequency pattern on the perception of duration. *J. Phonetics* 4: 113-117.

Leonard, R. J. and Ringel, R.L. (1979): Vocal shadowing under conditions of normal and altered laryngeal sensation. *Journal of Speech and Hearing Research* 22: 794-817.

Liberman, A.M. and Mattingly, I.G. (1985): The motor theory of speech perception revised. *Cognition* 21: 1-36.

Needham, J.G. (1934a): The time-error in comparison judgments. *The Psychological Bulletin* 31: 229-243.

Needham, J.G. (1934b): The time-error as a function of continued experimentation. *Amer. J. Psychol.* 46: 558-567.

Needham, J.G. (1935): Contrast effects in judgments of auditory intensities. *J. Exper. Psychology* 18: 214-226.

Ohala, J. (1972): How is pitch lowered? *J. Acoust. Soc. Am.* 52: 124(A).

Ohala, J. and Ewan, W.G. (1973): Speed of pitch change. *J. Acoust. Soc. Am.* 53: 345 (A).

Oppenrieder, W. (1989): Fokus, Fokusprojektion und ihre intonatorische Kennzeichnung. In: Altmann, H., Batliner, A. and Oppenrieder, W. (eds.): 267-280.

Peak, H. (1939): Time order error in successive judgments and in reflexes: I. Inhibition of the judgment and the reflex. *J. Exper. Psychol.* 25: 535-565.

Peak, H. (1940a): Time order error in successive judgments and in reflexes: II. As a function of the first stimulus of a pair. *J. Exper. Psychol.* 26: 103-115.

Peak, H. (1940b): Time order error in successive judgments and in reflexes: III. Time error theories. *The Psychol. Review* **47**: 1-20.

Pierrehumbert, J.B. (1980): *The Phonology and Phonetics of English Intonation.* Ph.D. diss., MIT.

Pike, E.V. (1948): Problems in Zapotec tone analysis. *Int. J. Am. Ling.* **14**: 161-170.

Pike, E.V. (1974): A multiple stress system versus a tone system. *Int. J. Am. Ling.* **40**: 169-175.

Repp, B. H. (1984): Categorical perception: Issues, methods, findings. In: Lass, N. J. (ed.): *Speech and Language.* Orlando, etc.: 243-335.

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.

Repp, B.H. and Crowder, R.G. (1990): Stimulus order effects in vowel discrimination. *J. Acoust. Soc. Am.* 88: 2080-2090.

Schiefer, L. - Batliner, A. (1988): Intonation, Ordnungeffekt und das Paradigma der Kategorialen Wahrnehmung. In: Altmann, H. (ed.): *Intonationsforschungen.* Tübingen: 273-291.

Smith, M.R. (1976): *An investigation of changes in categorical perception of vowels.* Unpubl. Manuskript. Department of Linguistics, University of Connecticut. (Quoted by Repp et al. 1979).

Stott, L.H. (1933): *The discrimination of short tonal duration.* Ph.D. diss. University of Illinois Library. (Quoted by Woodrow 1951).

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

Studdert-Kennedy and Hadding, K. (1973): Auditory and linguistic processes in the perception of intonation contours. *Language and Speech* **16**: 293-313.

Sundberg, J. (1979): Maximum speed of pitch changes in singers and untrained subjects. *J. Phonetics* **7**: 71-79.