THE FIRST STEP TO A BETTER UNDERSTANDING OF VOWEL REDUCTION

Dick R. van Bergem

1. INTRODUCTION

One of the problems in speech synthesis and automatic speech recognition is how to cope with the great variability in the realizations of vowel phonemes. In speech synthesis we do not know how to introduce this variability in a systematic way in order to increase intelligibility and naturalness, whereas in automatic speech recognition it causes a great deal of labeling problems. Part of this variability can be ascribed to a phenomenon known as vowel reduction: A shrinking of the vowel space in the direction of the neutral schwa (/ə/).

From an articulatory point of view vowel reduction may be explained as an effect caused by the limitations of the articulators. In this view the limiting factor is time. This means that a minimum vowel duration is needed to reach the theoretical 'target position'. Factors mentioned in the literature that might influence vowel reduction in this way are speech rate, stress and local context (neighbouring consonants).

On the other hand vowel reduction may also be explained by the effect that a talker in a normal speech situation 'deliberately' minimizes his articulatory movements in order to restrict articulatory effort. This means that the talker doesn't aim at reaching theoretical 'target positions', but that he only wants to make faint articulatory movements in the right direction. From a perceptual point of view this would be fully justified, as listeners have an abundance of clues other than the acoustic ones at their disposal which enable them to receive the talker's message also when his speech sounds are not clearly pronounced. The limiting factor in this view is the amount of acoustic clues a listener can do without: How sloppy can one talk and still be understood? Factors mentioned in the literature that might be related to vowel reduction in this way are of a socio-phonetic or linguistic nature: Is a person talking in a spontaneous way or is he reading a text (speech situation)? Is the talker a professional (trained) talker? Do the vowels appear in function words or non-function words? Do the vowels appear in familiar (high-frequency) words or unfamiliar (low-frequency) words?

For an extensive literature overview of the above mentioned factors see Koopmans-van Beinum (1980). The first aim of the present investigation was to search for an experimental design that would enable us to measure vowel reduction with as many controlled parameters as possible. The second aim was to find support for either of the visions mentioned above. We will discuss two experiments in which only effects on steady-state vowel parts have been studied. In chapter 2 we will explain in what way the steady-state part of a vowel was defined and how we measured the distance from a steady-state vowel part to the 'neutral' vowel position. In chapter 3 CVC nonsense words are studied in two conditions: in isolation and in a carrier sentence. In chapter 4 vowels extracted from more natural speech material are studied. In this experiment we used three conditions: isolated words, read texts and free conversation. Chapter 5 gives conclusions.

2. MEASURING PROCEDURES

2.1. Definition of the steady-state vowel part

The speech segments that had been recorded for our experiments were lowpass filtered at 4500 Hz and digitally stored in the computer at a rate of 10000 samples per second. Subsequently, all vowels that we wanted to analyze were segmented from the digital speech waveform. This was done by looking at changes in the structure of the speech signal that was displayed on a graphics screen. For some vowels, for instance those surrounded by plosives, a clear beginning and end could be seen. For other vowels, especially those surrounded by nasals, liquidae or glides, this was much more difficult. Replaying selected parts of the speech signal in order to locate the vowel boundaries in these difficult cases by ear didn't help much either, because the influence of nasals, liquidae and glides can often be heard throughout the entire vowel. Therefore we decided to use only the visual display of the speech signal to segment the vowels, which was done with the greatest possible consistency. After the segmentation of the vowels their formant frequencies were measured with a 25 ms. Hamming window that was shifted each ms. and a 10th-order LP-analysis that makes use of the Split-Levinson algorithm (see Willems, 1986).

Usually the position of the steady-state vowel part is determined by the experimenter who makes use of phonetic knowledge. Because it is not always clear which (subjective) criteria are used by the experimenter and whether the criteria are applied consistently, we preferred to design a computer algorithm that could trace the steady-state vowel part. This algorithm is based on an idealized vowel concept in which the formant tracks show the pattern: transition - steady-state part - transition (see figure 1). This means that the formant frequencies are shifting in the beginning (transition from the preceding phoneme), then remain constant for some time (the steady-state part) and finally shift again (transition to the following phoneme). Using this model we can make the following assumptions for the steady-state part in actual vowel realizations:

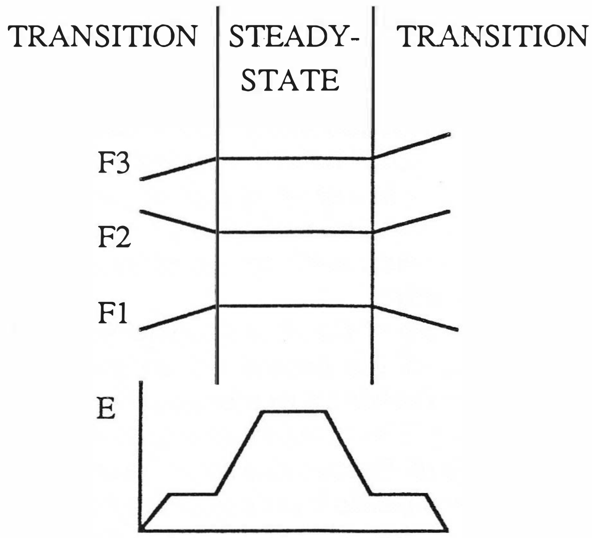

Figure 1. Formant tracks of an ideal vowel. Each formant contains two transitions and a steady-state part. At the bottom the spread of energy in the vowel is shown.

1. The formant tracks of F1, F2 and F3 must be as straight as possible in the steady-state vowel part.
2. The steady-state vowel part is more likely to occur in the middle of the vowel than at the edges.
3. In addition, steady-state vowel parts often contain much energy, so: The steady-state vowel part is more likely to occur at a place with a relatively high energy.

The first criterium was met by moving a window through the formant tracks and by measuring the within variance of each formant track inside the window. In order to give less weight to differences in higher frequencies, all formant frequencies in Hz were transformed into logarithmic values. The steady-state vowel part was defined as that part where the pooled within-variance inside the window is minimal. The steady-state formant values were obtained by averaging the (log-transformed) formant values within the selected vowel part. The window size should not be too small in order to prevent the detection of local steady-state vowel parts. On the other hand a very broad window can only make a coarse estimation of the steady-state vowel formants due to the averaging operation. A (time normalized) window size of 1/4 of the vowel length proved to be a reasonable compromise.
The second and third criterium were met by using two weighting factors. The first weighting factor was based on the position of the window within the vowel. A window at the middle of the vowel was given a weighting factor of 4 and a window at either edge of the vowel a weighting factor of 1. Weighting factors for window positions between these extremes were obtained through parabolic interpolation.The second weighting factor was the rms-value of the speech signal samples inside the window. The weighting factors that were chosen on a trial-and-error basis were introduced to avoid for instance the selection of steady-state vowel parts that occurred sometimes at the end of a vowel followed by the consonant /r/. In most cases the weighting factors did not influence the choice of the steady-state vowel part, because the pooled within variance (for three formants) differed considerably between several windows. Only when the variance inside several windows was of the same order, a window position near the edge of the vowel and/or with a relatively low energy would be rejected in favour of one near the middle of the vowel containing more energy.
A computer program was written that displayed formant tracks and the spread of energy within a vowel together with the position of the steady-state vowel part that had been chosen by the algorithm. All vowels that had been selected for our experiments were carefully examined in this way. The choice of the position of the steady-state vowel parts by the algorithm was almost always very satisfactory. In a few cases the choice was doubtful: when the tracks were very unstable, when the tracks reached steady-state positions at different moments or when the vowels were slightly diphthongized (mainly /o/, /e/, /ϕ/). However, in none of these cases arguments could be found to change the chosen position of the steady-state vowel part, so none of the positions that had been chosen by the algorithm were manually corrected.

2.2. Definition of the distance measure

Now that we have defined the way to obtain the formant frequencies of the steady-state vowel part we also have to define the 'neutral' vowel position. Following Koopmans-van Beinum (1980) we calculate a vowel centroid which is composed of the average values of all (log-transformed) first, second and third formant frequencies of a talker's monophthongs. The distance measure we chose was the Euclidian distance between the steady-state vowel and the centroid in a three dimensional formant space:

$$d=\{ (\log F1_V-\log F1_C)^2 + (\log F2_V-\log F2_C)^2 + (\log F3_V-\log F3_C)^2 \}^{1/2}$$

in which d means distance, Fi_V (i=1,3) means frequency of formant i of (the steady-state part of) a *vowel* and Fi_C (i=1,3) means frequency of formant i of the *centroid*.

3. FIRST EXPERIMENT

3.1. Design

In order to study the effect of neighbouring consonants on the steady-state vowel part, all Dutch monophthongs /α, ɔ, ɛ, I, œ, u, i, y, a, o, e, ɸ/ were spoken by a male speaker (the author) in a symmetric CVC context (C=/p, t, k, f, s, χ, r, l, m, n, j, w/). All recordings were made in an anechoic room with a Sennheiser MD421N microphone and a Revox A77 tape recorder. In the first condition the CVC words were pronounced in isolation; in the second condition they were embedded in the carrier sentence "Nu krijgt de CVC een beurt" ("Now gets the CVC a turn"). By using this Dutch carrier sentence the CVC words are surrounded by the neutral schwa, so that the carrier sentence doesn't have a great coarticulatory effect on the test words. The sentence accent was placed on the word "beurt" and consequently the test word had no major stress. The total number of vowels was 288 (12 vowels x 12 consonants x 2 conditions). The talker's centroid was obtained by averaging the steady-state formant frequencies of all the vowels from both conditions.
The aim of this experiment was twofold. In the first place we wanted to establish how much the steady-state vowel part is influenced by the consonants. In the second place we wanted to establish whether the vowels of the test words in the carrier sentence that were spoken with less stress were more reduced than those of the well pronounced stressed test words that were spoken in isolation.

3.2. Results

In the upper part of figure 2 the logarithmic values of the first and second formant of the steady-state part of all vowels from condition 1 (test words spoken in isolation) have been plotted. In the lower part of figure 2 the same has been done for the vowels from condition 2 (test words in the carrier sentence).
In figure 3 both conditions are compared in a direct way by plotting the logarithmic F1-values from the condition isolated words on the abscissa and the logarithmic F1-values of the corresponding vowels (occurring in the same word) from the condition carrier sentence on the ordinate. In the lower part of figure 3 the same has been done for F2-values. Vowel tokens on the diagonal of the figures indicate a perfect match of the F1-values or F2-values from both conditions. Vowel reduction would be indicated by a majority of vowel tokens above the diagonal in the bottom part of each figure and a majority of vowel tokens below the diagonal in the upper part of each figure. It can be seen that the vowel tokens are evenly spread around the diagonal both for F1-values and F2-values indicating that there is no reduction effect. This is confirmed by the regression lines (the dashed lines) that have been drawn in the figures: These lines almost coincide with the diagonal. Correlation coefficients are 0.92 for F1-values and 0.97 for F2-values. Coarticulation is indicated by the spread of the 12 tokens per vowel along the regression lines.

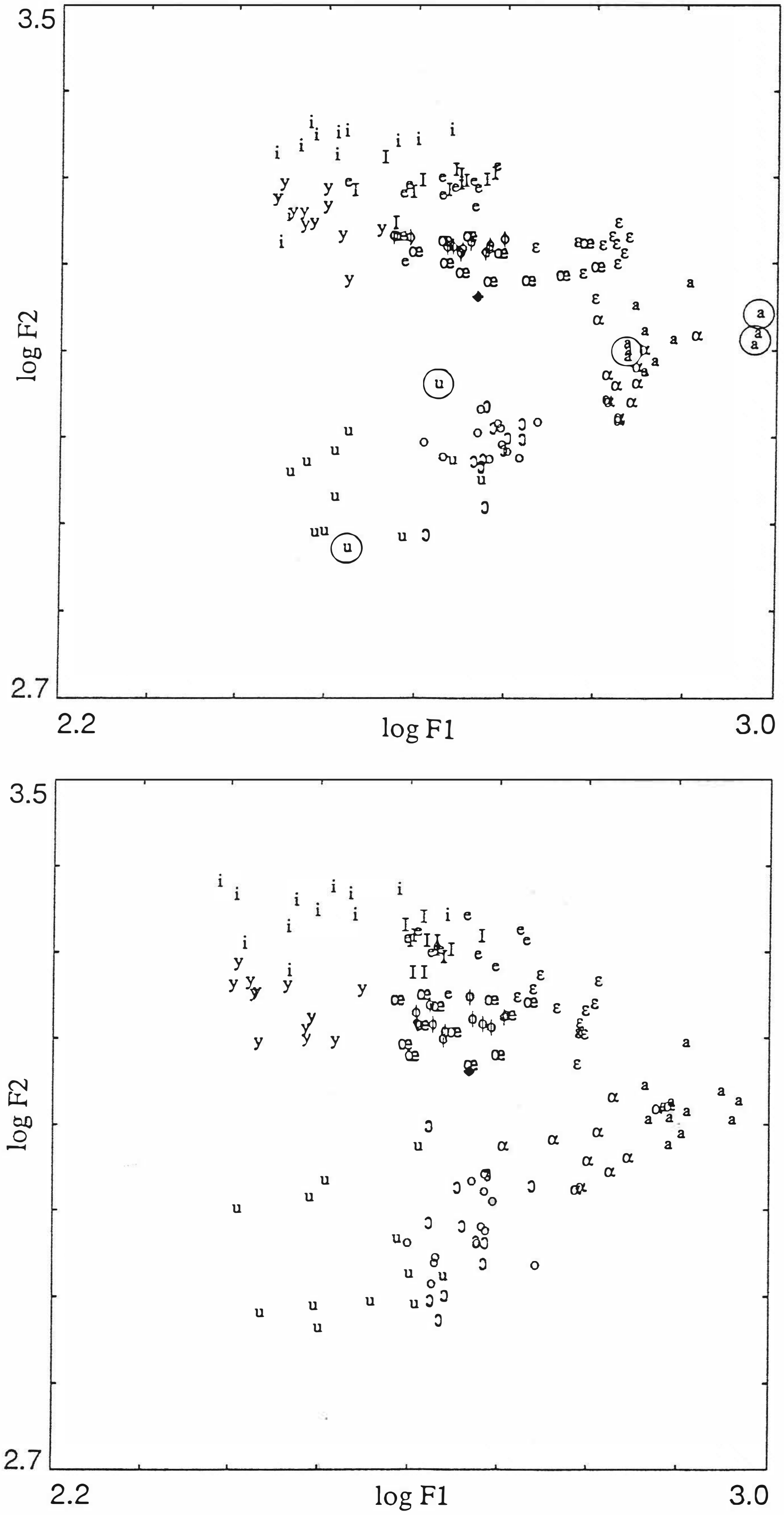

Figure 2. Logarithmic formant plots of 12 monophthongs in two conditions: isolated CVC words (upper part) and CVC words in a carrier sentence (lower part). The centroid is indicated by a dark square. Tokens of the vowel /a/ surrounded by plosives and fricatives have been encircled as well as the /u/ from the word /tut/ and the word /χuχ/, see text.

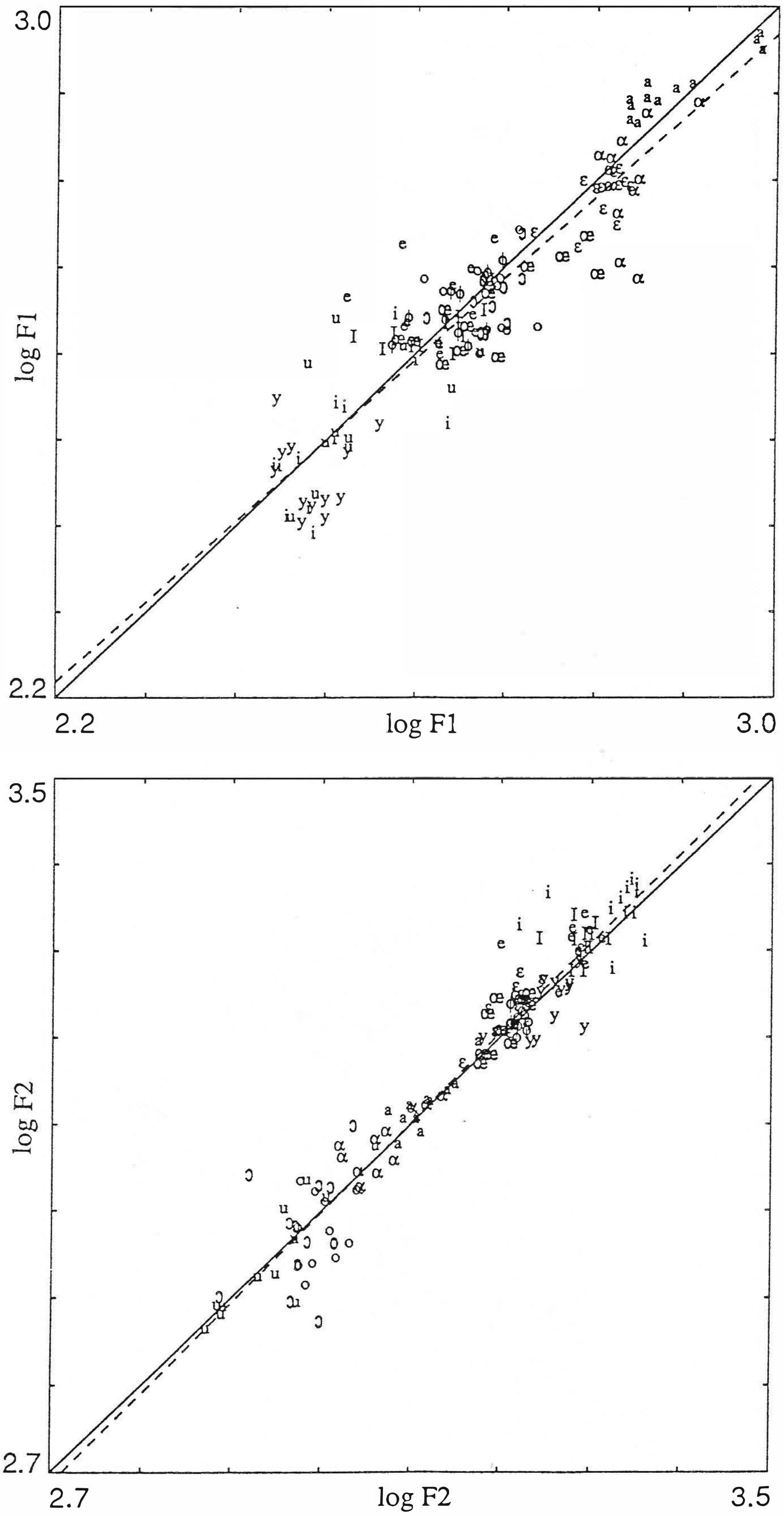

Figure 3. Comparison of logarithmic F1 values (upper plot) of corresponding vowels in the condition isolated words on the abscissa and the condition carrier sentence on the ordinate and the same comparison of logarithmic F2 values (lower plot). Dashed lines are regression lines.

Another way to study vowel reduction is to look at the distance of each vowel to the centroid. If the vowels in sentence context were more reduced than those spoken in isolation we would expect the distances to the centroid in condition 2 to be shorter than those in condition 1. However, a two tailed t-test revealed no significant difference between distances in condition 1 and condition 2. Another t-test was carried out on vowel durations in condition 1 and condition 2. This test revealed that vowel durations in condition 2 were significantly smaller than those in condition 1 ($t=5.23$, $p<0.01$). Average formant frequencies and durations together with their standard deviations for each vowel are given in table 1. For the sake of clarity all logarithmic formant values have been retransformed to frequencies in Hz which are much easier to compare. Average values and standard deviations were obtained with formant frequencies in Hz.

Table 1. Mean formant frequencies (in Hz) and durations (in ms.) with their standard deviations for each vowel in each condition.

	isolated words						carrier sentence					
	F1	sd	F2	sd	dur.	sd	F1	sd	F2	sd	dur.	sd
ɑ	690	49	1176	102	124	68	643	88	1196	102	90	36
ɔ	481	31	951	86	123	57	459	38	954	148	92	44
ɛ	647	42	1647	87	114	46	609	40	1716	117	81	27
ɪ	426	46	1975	82	108	49	427	24	2046	91	82	38
œ	489	84	1607	74	110	52	450	53	1646	109	82	32
u	357	61	902	120	121	64	352	59	904	150	90	40
i	333	51	2146	195	115	72	321	57	2246	150	79	22
y	310	26	1822	121	123	69	293	32	1755	128	95	34
a	789	111	1321	90	183	64	820	66	1336	96	149	29
o	486	40	1004	45	153	25	469	44	948	84	132	22
e	427	47	1929	121	153	23	462	49	2041	112	125	15
ϕ	451	38	1677	26	162	25	454	36	1677	53	136	18
average	491		1513		132		480		1539		103	

The limited data set didn't allow us to reliably investigate the specific effects of particular consonants. Nevertheless we will give two examples of coarticulation effects, one for the vowel /a/ and one for the vowel /u/, that seem to be consistent, because the effects occurred in both conditions.

1. The (steady-state) first formant of the vowel /a/ is lower when surrounded by one of the fricatives /f, s, χ/ than when it is surrounded by one of the plosives /p, t, k/ (see table 2 and also figure 2). This effect is demonstrated in the upper part of figure 4 for the words /tat/ and /sas/ from the condition carrier sentence. Apparently the steady-state /a/ is produced with a higher jaw when it is surrounded by fricatives than when it is surrounded by plosives, because the position of the jaw is reflected in the first formant frequencies (Pickett, 1980). An explanation for this effect might be that the constriction that is needed to produce fricatives is counteracting the lowering of the jaw, especially with labial and alveolar fricatives.
2. The same effect is demonstrated for the vowel /u/ with the words /tut/ and /χuχ/ from condition 2 in the lower part of figure 4. In this case the F1 of the /u/ is also somewhat higher when it is surrounded by plosives than by fricatives (table 2). In

addition to the differences in F1 for the vowels in the words /tut/ and /χuχ/, there is also a striking difference in F2 (see table 2 and also figure 2). According to Pickett (1980) the frequency of a vowel's F2 is raised by a front tongue constriction. The latter difference might thus be explained by the fact that the tongue is moved to the front in order to make a constriction for the /t/ with the tongue tip at the alveolus, whereas the tongue position for the production of the vowel /u/ is at the back of the mouth. In this case the counteracting movements result in a very high second formant of the /u/. The second formant of the /u/ was also relatively high (although less high than with /tut/) for the neighbouring consonants /s, l, n, j/.

Table 2. First and second formant frequencies (in Hz) of the vowel /a/ and the vowel /u/ for six neighbouring consonants in both conditions.

	a				u			
	F1		F2		F1		F2	
	word	sentence	word	sentence	word	sentence	word	sentence
p-p	956	923	1290	1285	442	365	948	791
t-t	965	939	1329	1351	426	412	1160	1198
k-k	970	898	1400	1387	476	403	899	852
f-f	695	775	1268	1331	311	275	783	766
s-s	694	747	1248	1288	292	259	917	1013
χ-χ	725	791	1338	1347	338	320	752	738

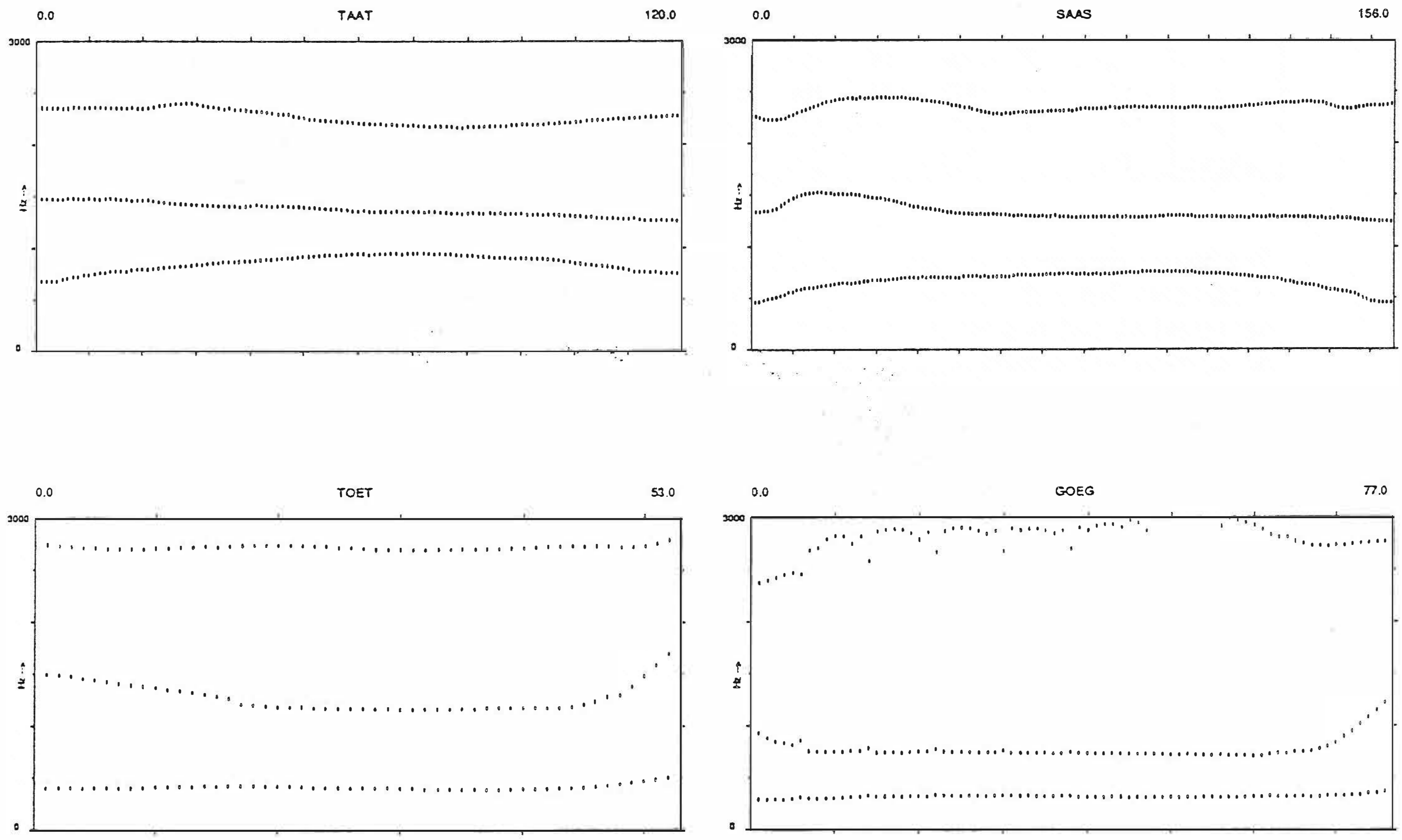

Figure 4. Formant tracks of the vowels in the words /tat/, /sas/, /tut/ and /χuχ/ from the condition carrier sentence

3.3. Discussion

We may ask ourselves why the first formant frequency of the vowel /a/ in /sas/ is lower than the first formant frequency of the vowel /a/ in /tat/ and why the second formant frequency of the vowel /u/ in /tut is higher than the second formant frequency of the vowel /u/ in /χuχ/ (table 2). One reason might be the limiting factor of *time*. That is, the articulatory movement to the vowel target cannot be completed, because the transition from the vowel to the next phoneme has to start before this has happened. If this were true the formant tracks of the vowel would only consist of (unfinished) transitions. However, figure 4 shows that the F1 in the word /sas/ and the F2 in the word /tut/ have a considerable steady-state part. It looks as if we (as speakers) set bounds to the 'distance' we have to bridge from one articulatory position to another, in order to restrict articulatory effort. If the 'distance' is too great we only cover part of it which results in a less pronounced vowel.

4. SECOND EXPERIMENT

4.1. Design

Apart from coarticulatory effects, we didn't find a reduction effect caused by stress in the experimental design of the previous chapter. Apparently a limited decrease in stress doesn't cause vowel reduction. Recently the effect of speech rate on vowel reduction has been investigated by van Son and Pols (1988). They compared the vowels from a text read aloud by a trained speaker in two conditions. In one condition the text was read at a normal speed and in the other it was read at a fast speed. They didn't find significant differences between the 'stationary' formant frequencies of the vowels in the two conditions, although the speaker's vowels were significantly shorter in the fast rate condition.
Since vowel reduction is most likely to occur in more spontaneous speech (Koopmans-van Beinum, 1980), we decided to compare the vowels from free conversation with those from a read text and with words spoken in isolation. In order to keep coarticulatory effects in control, the following design was chosen. First a male speaker (different from the first one) was invited to talk with a free word choice about his work, his holidays, his favourite meal etc. Subsequently part of his talking (about his favourite meal) was written down. This text was cleaned up for reading with regard to repetitions and deletions, but we attempted to change the original words as little as possible. Next the same speaker was asked to read this text aloud as well as a list of (mono- and multi-syllabic) words that were selected from it. In this way each selected word occurred in three different conditions: in isolation, in a read text and in free conversation. From the set of monophthongs the vowels /i, a, ɔ / had been selected for investigation, because these vowels have a reasonable frequency of occurrence, do not diphthongize and represent more or less the extremes of the vowel triangle. In the written texts we searched for all words containing one of these vowels, but these words were only selected if the consonantal context of the vowel was different from the earlier selected words with the same vowel. If a word began or ended with one of the vowels that we wanted to analyze the preceding or following word in the text was included in the condition isolated words in order to preserve coarticulatory effects. In such a case the 'isolated word' consisted of two words. For each vowel 30 words were selected resulting in a total of 270 words (3 vowels x 30 words x 3 conditions). Apart from the speech material mentioned above we had also recorded several series of our talker's monophthongs spoken in isolation. Since we only studied 3 vowels, the centroid was

obtained by averaging the formant frequencies of all monophthongs that had been spoken in isolation by our talker. All recordings were made in an anechoic room with a Sennheiser MD421N microphone and a Revox A77 tape recorder.

4.2. Results

The first thing to mention is the fact that some of the vowels that had been selected from the written texts were missing in the actual speech material: three specimens of the vowel /a/ in free conversation, two specimens of the vowel /i/ in free conversation and one in the read text, two specimens of the vowel /ɔ/ in free conversation. The eight missing vowels were four times embedded in function words (daar, die, nog, of) and four times in non-function words (vakantie, natuurlijk, citroenzuur). The number of missing vowels is too small to draw conclusions, but it should be mentioned that most of these vowels were rather short and were relatively close to the centroid in the other conditions. For statistical analyses the words in which these vowels occurred were omitted in all conditions.

In figure 5 the logarithmic values of the first and second formant of all 3 x 30 vowels in three conditions have been plotted. The figure at the top left shows the vowels spoken in isolation (average of two series) by our talker. This figure has been added as a reference frame. The figure at the top right shows the vowels of the isolated word condition, the figure at the bottom left shows the vowels of the read text condition and the figure at the bottom right the vowels of the free conversation condition.

Table 3 gives the mean formant frequencies and the mean vowel durations together with their standard deviations for each vowel in each condition (All logarithmic formant values were transformed to frequencies in Hz. All statistics were done on frequencies in Hz). A one-way analysis of variance on the distances between vowels and centroid in the three conditions reveals a significant effect for the vowel /i/ ($F = 27.16$, $p<0.01$) and for the vowel /ɔ/ ($F = 24.15$, $p<0.01$), but no effect for the vowel /a/ ($F = 0.75$, $p>0.52$). Tukey's HSD tests show that for the vowel /i/ all conditions are significantly different from each other and that for the vowel /ɔ/ only the condition free conversation is significantly different from the other two conditions. A one-way analysis of variance on vowel durations in three conditions shows a significant effect for all vowels: /i/ ($F = 8.04$, $p<0.01$), /ɔ/ ($F = 10.12$, $p<0.01$) and /a/ ($F = 17.73$, $p<0.01$). Tukey's HSD tests reveal that only the vowel duration in the condition isolated words is significantly different from the other two conditions for each of the three vowels.

Table 3. Mean formant frequencies (in Hz) and durations (in ms.) with their standard deviations for each vowel in each condition. Standard deviations for formant frequencies and durations of isolated vowels are meaningless (only two vowels). These values have been omitted.

	isolated vowels			isolated words			read text			free conversation		
	/a/	/i/	/ɔ/	/a/	/i/	/ɔ/	/a/	/i/	/ɔ/	/a/	/i/	/ɔ/
F1	768	259	489	587	243	419	567	265	381	603	299	417
sd F1	-	-	-	54	28	43	88	29	51	67	28	35
F2	1232	2125	887	1203	2040	837	1219	1968	921	1223	1949	1031
sd F2	-	-	-	73	88	92	97	93	112	92	89	135
F3	2512	2994	2663	2536	2754	2560	2472	2586	2517	2510	2537	2495
sd F3	-	-	-	205	165	253	185	161	240	129	131	187
dur.	295	193	193	192	97	102	121	65	71	127	69	73
sd dur.	-	-	-	52	39	31	42	23	26	51	34	29

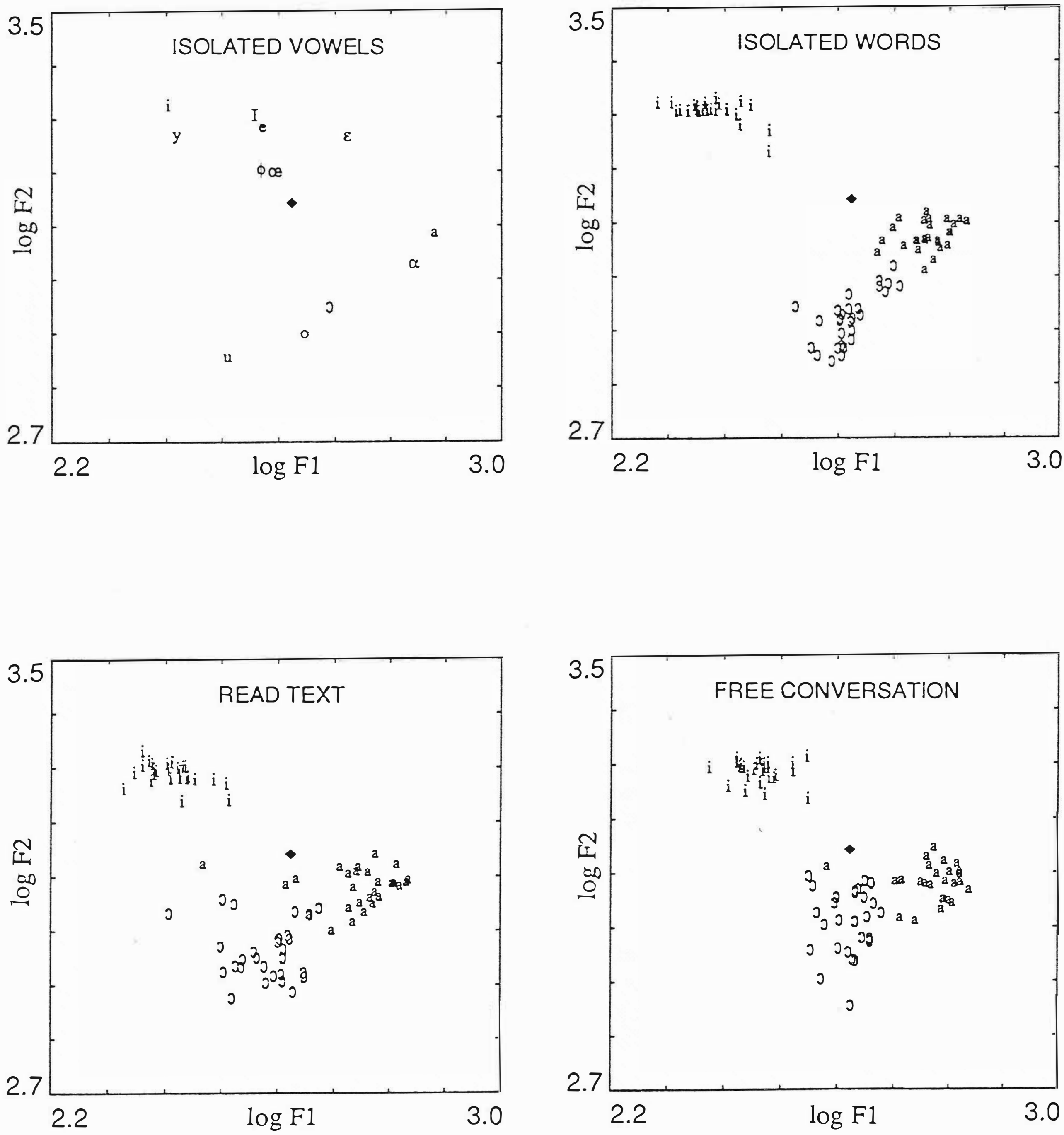

Figure 5. Logarithmic formant plots of the vowels /a/, /i/ and /ɔ/ in four different conditions. The conditions are indicated in the plots. The centroid is indicated by the dark square.

In figure 6 two plots are shown. The left plot gives the (logarithmic) F1-values of all vowels from the condition isolated words on the abscissa and the F1-values of the corresponding vowels form the condition free conversation on the ordinate. In the right plot the same has been done for corresponding F2-values. These figures reveal that both F1- and F2-values of the vowel /i/ are reduced in the condition free conversation: Almost all vowel tokens of the /i/ are above the diagonal of the F1-plot (nearer to the F1 of the /ə/) and below the diagonal of the F2-plot (nearer to the F2 of the /ə/; although this is not very clear from the F2-plot 24 /i/ tokens on a total of 27 are below the diagonal). For the vowel /ɔ/ a clear reduction effect can be seen in the F2-plot. Vowel tokens of the vowel /a/ are evenly spread around the diagonal both in the F1-plot and the F2-plot.

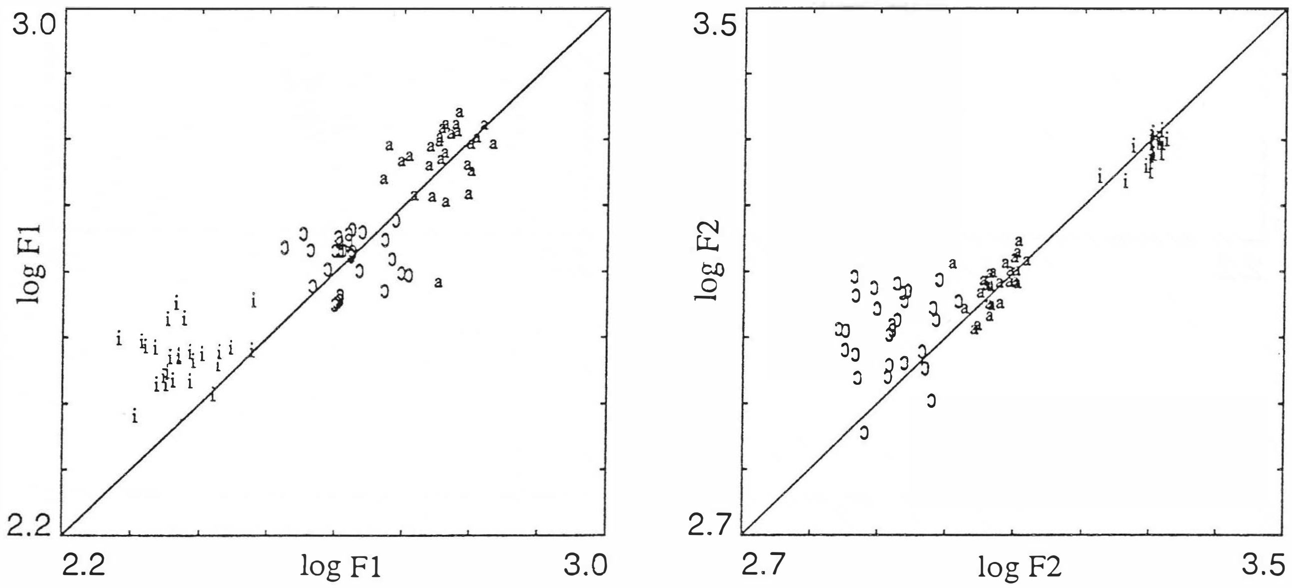

Figure 6. In the plot on the left a comparison has been made of logarithmic F1 values of corresponding vowels in the condition isolated words on the abscissa and free conversation on the ordinate. In the plot on the right the same comparison has been made of logarithmic F2 values.

4.3. Discussion

How can we explain the reduction effect that occurs in some cases and how can we explain the lack of a reduction effect in the other cases? The factor stress might have played an important role in this experiment. The relation between stress and reduction in our experiment can be formulated in two hypotheses:

1. The amount of stress or the number of stressed vowels decreases in more spontaneous speech situations, which results in vowel reduction.
2. Within one condition stressed vowels are less reduced than unstressed vowels.

In order to test these hypotheses we used vowel duration as one of the acoustic correlates of stress. The first hypothesis then predicts that vowel durations are shorter in conditions with more reduced vowels. This effect, however, is not confirmed by our data; the vowel durations between the conditions read text and free conversation were not significantly different for the vowel /i/ and the vowel /ɔ/, but we did find a significant difference in distances to the centroid between these conditions. On the other hand we did not find a significant distance effect for the vowel /a/, but this vowel was significantly longer in the condition isolated words. The second hypothesis predicts a positive correlation between distance to the centroid and vowel duration within one condition. These correlations are given in table 4. It will be clear that the weak correlations do not support the second hypothesis either.

Table 4. Correlation between duration and distance to the centroid for each condition.

	isolated words	read text	free conversation
/a/	0.28	0.22	0.48
/i/	0.26	0.30	0.69
/ɔ/	-0.12	-0.12	0.24

Since duration will probably not be an optimal indicator of stress, we asked a phonetically trained listener to establish which vowels were stressed and which were not. The main criterion for stress was word stress, but vowels with word stress were nevertheless marked as unstressed if stress was absent due to the sentence accent pattern. The distinction between primary and secundary accent was thus not used; a vowel not bearing a primary accent was defined as unstressed. Table 5 gives the number of stressed and unstressed vowels in each condition.

Table 5. The number of stressed and unstressed vowels in each condition according to a phonetically trained listener.

	isolated words		read text		free conversation	
	stressed	unstressed	stressed	unstressed	stressed	unstressed
/a/	20	7	10	17	16	11
/i/	12	15	12	15	11	16
/ɔ/	21	7	11	17	13	15

The first hypothesis predicts that less vowels are stressed in the conditions with more reduced vowels. Table 5 shows that the number of stressed vowels is different in the three conditions for the vowel /a/, but this vowel didn't reveal reduction effects. On the other hand, the number of stressed vowels is the same in the three conditions for the vowel /i/, but this vowel did have a significant reduction effect. Unfortunately the second hypothesis could not be tested in this way, because the number of unstressed vowels was too small in some cases (see table 5).

Nevertheless it will be clear that under the present experimental conditions stress cannot be seen as an important reduction factor. Since we used the same speaker and the same words in all three conditions, the reduction effects that occurred can only be ascribed to the factor 'speech situation' itself. Apparently our talker speaks with a minimal articulatory effort in the condition free conversation, which makes his vowel space small in comparison with for instance his vowels spoken in isolation (see table 3 for a comparison of formant frequencies). However, a listener has an abundance of information sources other than acoustic ones at his disposal, which allow him to receive the talker's message with great ease even if part of the acoustic clues are ambiguous or even missing. In some cases vowels were not spoken at all by our talker (see paragraph 4.2), but listeners are very well able to fill in these gaps and may not even notice the omissions. For the condition read text our talker increased his articulatory effort, which is quite common when people read aloud. This can be seen as an 'exaggeration' of the articulatory movements which were even greater when only isolated words were read

aloud. This 'exaggeration' has blown up the vowel triangle. Since free conversation is in my view 'normal' speech and read texts or read words are more artificial speech styles, it seems more appropriate to talk of vowel *expansion* than of vowel reduction. For the vowel /a/ our talker apparently didn't 'exaggerate' the lowering of his jaw in the more artificial speech styles, so that the first formant of this vowel didn't significantly change which explains the lack of a reduction (expansion) effect for this vowel.
The present experimental design seems well fit to investigate vowel reduction, especially because vowels in identical words can be studied in different experimental conditions including free conversation. For future research we will therefore first extend the small data set we already have with the same speaker. This will give us the opportunity to investigate the possible influence of stress on vowel reduction more thoroughly.
In the second place it will be interesting to look at the effect of vowel reduction on the dynamic properties of the vowels: In what way do formant transitions of vowels from different conditions (e.g. isolated words versus free conversation) change when the formant frequencies of the steady-state part are closer to the neutral schwa? For this purpose we plan to use the same simple vowel model that was discussed earlier in this article.
In view of the findings in the present investigation it may also be worthwile to look some more at other than acoustic reduction factors. Thus we might pay some attention to linguistic factors such as word class or word frequency in relation with vowel reduction.

5. CONCLUSIONS

- Under the experimental conditions of the first experiment coarticulation had a considerable effect on steady-state vowel parts. In the same experiment an effect of stress with respect to vowel reduction could not be demonstrated.

- The experimental design of the second experiment, starting from free conversation and from there choosing the words in the other conditions accordingly, seems to be very well fit to study the effect of vowel reduction in a detailed way.

- Under the experimental conditions of the second experiment vowel reduction was found for only a part of the vowels. No clear relation between stress and vowel reduction could be demonstrated. It appears that the factor speech situation (speech style) plays an important role in vowel reduction. This may also imply that socio-phonetic factors (and maybe also linguistic factors) have a greater effect on vowel reduction than acoustic-phonetic factors.

ACKNOWLEDGEMENTS

I am grateful to Rob van Son for his assistance in plotting the figures and to Louis Pols and Florien Koopmans-van Beinum for their comments on earlier versions of this article.

REFERENCES

Koopmans-van Beinum, F.J. (1980), Vowel Contrast Reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions, Doct. diss., University of Amsterdam.

Pickett, J.M. (1980), The sounds of speech communication. A primer of acoustic phonetics and speech perception, University Park Press, Baltimore.

Son, R.J.J.H. van & Pols, L.C.W. (1988), 'Differences in formant values of Dutch vowels due to speaking rate', Proceedings of Speech '88, 7th FASE Symposium, Edinburgh, Book 1, 313-320.

Willems, L.F. (1986), 'Robust Formant Analysis', IPO annual progress report 21, 34-40.