PERCEPTION OF THE LONG - SHORT VOWEL OPPOSITIONS IN DUTCH: A GATING STUDY

Laurens W.M. Damen

INTRODUCTION

This paper describes an experiment belonging to a research project titled *The effect of local context for the correct interpretation of acoustic-phonetic information in speech understanding*. The subject of the project is perceptual normalization due to the rate of speech of the acoustic-phonetic context. Its purpose is to investigate tentatively to what extent the long-short vowel oppostion pairs in Dutch are perceived in relation to the speech rate of the local acoustic-phonetic context.
The primary cue to differentiate perceptually between the two vowels of a long-short vowel opposition pair is traditionally thought to be the length or duration of the vowels, their spectral contents being only minimally different. If this view is correct and if speech is perceived in a context dependent manner, it must be possible to influence the perception of the long vowels by the speech rate of the acoustic-phonetic context.
Some experiments of this kind were carried out by Johnson and Strange (1982). They asked subjects to identify vowels in tVt syllables that were spoken in a fast or in a normal speech tempo. The vowels of the normal-rate spoken syllables were always identified correctly, whether they were presented in a rate-inappropriate sentence frame or in their original sentence frame. The fast spoken syllables, however, were identified in a speech rate dependent manner. Significantly more vowel identification errors were made, when the fast spoken syllables were embedded in a normal-rate spoken sentence frame than when presented in their appropriate fast spoken sentence frame. Analyses of the errors revealed that the majority of errors consisted of confusions between long vowels and their spectrally similar short counterparts.
Van Bergem and Drullman (1985) repeated one of the experiments of Johnson and Strange (1982) for the Dutch language, but did not obtain comparable results. In their experiment, almost all vowels were identified correctly, irrespective of the speech rate of both the tVt syllables and the acoustic-phonetic context. This failure to find any influence of the contextual speech rate on vowel identification is possibly due to the fact that the high percentage correct vowel identifications totally obscured all effects of context.
Johnson and Strange (1982) as well as Van Bergem and Drullman (1985) used a simple phoneme identification task in their respective experiments. This task provides only information about the result of the process of vowel identification. In studying perceptual normalization in relation to the speech rate of an utterance, one clearly wants and probably needs a task that provides detailed information about what analyses the listeners performs on the speech input as it accumulates over time. The gating task developed by Grosjean (1980) is one that provides exactly the information needed.
Two aims underlie the present study. The first is to investigate the suitability of a gating task for the study of long-short vowel opposition pairs. The second aim is to compare the standard version of the gating task to a version of which portions of the words are replaced by envelope-shaped noise instead of deleted as in the standard version. The noise provides acoustical information about amplitude changes within the word as well

as information about the original duration of the gated word. On the one hand this information might enhance vowel and word recognition if speech and noise are integrated by the listener to a single percept. On the other hand, it may be possible that the noise masks speech to some extent and thereby delays the recognition of vowel and word.

METHOD

Materials

Experimental words The Dutch language has four long-short vowel oppositions, i.e. /a/ - /ɑ/, /o/ - /ɔ/, /e/ - /ɪ/ and /ɸ/ - /œ/. Twelve pairs of monosyllabic CVC (the initial and final C's can form consonant clusters as well) nouns were selected for each vowel opposition (e.g. /tak/ - /tɑk/, /bot/ - /bɔt/ , /strep/ - /strɪp/ and /pɸl/ - /pœl/). Most of the word pairs differed only with respect to the vowels, whereas the final consonant (cluster) of some of the noun pairs of the vowel opposition /ɸ/ - /œ/ varied too (e.g. /bɸk/ - /bœks/). Syllables and particularly vowels at the end of a sentence or clause tend to be lengthened; therefore, the words were uttered as the last word of a sentence. The sentences were nonsensical ones, comparable to those used by Nakatani and Dukes (1973).
Filler words In order to prevent the subjects, as well as possible, from recognizing the purpose of the experiment, 60 bisyllabic and 20 monosyllabic nouns were chosen to serve as filler words. None of these words contained any of the vowels used in the experimental words. They were spoken as the last word of a neutral carrier sentence.
Preparation of the stimuli All the sentences were read by a male speaker and recorded on audio tape. The recordings were made in an anechoic room with a Sennheiser MD412N microphone and a Revox A77 2-track tape recorder, using an Agfa PEM268 tape at a recording speed of 7.5 inch/sec. The recorded sentences were analog low-pass filtered at 4.5 kHz, digitized at a sampling rate of 10 kHz and stored on computer disk. The words were then excised from the digitally stored sentences with the aid of a digital waveform editor, by visually inspecting the speech wave and by using auditory feedback.

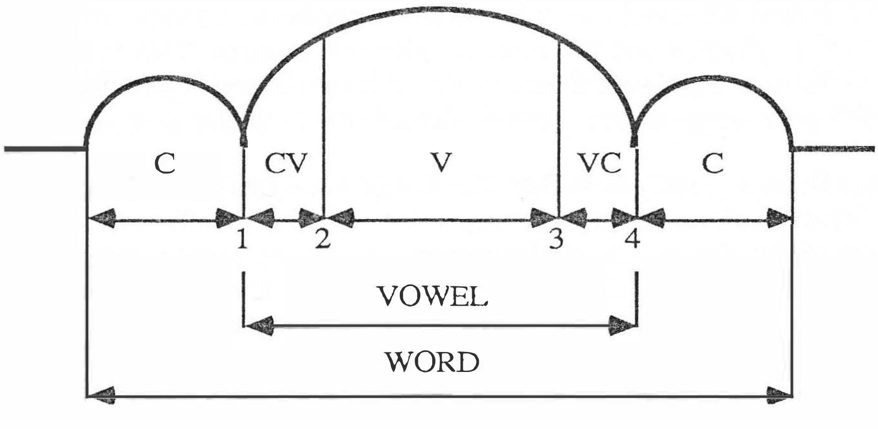

Figure 1. Schematic representation of a word. C = consonant, CV = the initial vocalic transition, V = the steady-state part of the vowel and VC = the final vocalic transition The vertical lines 1 to 4 are segment boundary markers.

In each experimental word markers were placed that defined the beginning (marker 1 in Figure 1) and the ending (marker 4 in Figure 1) of the vowel. Two additional markers were placed within the vowel. These markers defined the beginning (marker 2 in Figure 1) and the end (marker 3 in Figure 1) of the so-called steady-state part of the vowel. Thus, each experimental word was divided into five different speech segments: (1) the initial consonant (cluster) C; (2) the initial vocalic (CV) transition; (3) the steady-state part (V) of the vowel; (4) the final vocalic (VC) transition; and (5) the final consonant (cluster) C. The duration of the three vocalic (CV, V and VC) segments of each word have been measured, using the aforementioned waveform editor. See Table 1.
The stimulus set of each gated word was phonetically motivated. It was prepared in relation to the defined five different speech segments in a word. The first gate of a word consisted only of the initial consonant (cluster). The second gate contained the same speech segment plus the CV transition. From then on, gates were increased by two vocalic periods of the steady-state part of the vowel. Whenever the number of vowel periods in the steady-state part was not a multiple of two, the gate was increased by the appropriate amount of vowel signal upto the beginning of the final vocalic (VC) transition. In the penultimate gate of a word the complete vowel, including the VC transition was presented. Finally, the last gate of a word consisted of the entire, original stimulus word.
The stimulus sets of the filler words were prepared in a different manner. No attempt was made to divide the filler words into phonetically meaningful segments as was done with the experimental words. On all occasions, the first gate of a monosyllabic filler word lasted at least 100 msec and each subsequent gate was increased by 40 msec or more. The first gate of a bisyllabic filler word was at least 80 msec. Subsequent gates were once again always increased by a minimum of 40 msec.
Two sets of gated stimuli were generated for each word. In the Silence condition (shortened to S-condition) the appropriate part of a gated word was deleted. Most gating studies, reported in the literature, have used this procedure to create gates (e.g. Cotton & Grosjean, 1984; Grosjean, 1980, 1985; Tyler, 1984; Tyler & Wessels, 1983, 1985; Warren & Marslen-Wilson, 1987). In the other condition, the Noise-Replaced one (NR-condition), the segment of a word that was deleted in the S-condition, was replaced by noise with a long term average speech-like spectrum (speech noise). A similar procedure has recently been used by Salasoo and Pisoni (1985).

Design

A great many of the studies that employed the gating paradigm made use of the standard or sequential presentation format (e.g. Grosjean, 1980, 1985; Tyler & Wessels, 1983; Warren & Marslen-Wilson, 1987). In this presentation format, listeners hear successively larger fragments of a target word with each fragment increasing by some constant or variable amount. The fragments of a word are presented after each other until the whole word has been introduced. This sequential presentation format is not considered an appropriate one for the present experiment. By applying this format, listeners can hardly fail to associate a correct perception of the target vowels with the duration of the presented fragment. This knowlegde may influence the responses of the listeners in unpredictable ways. To avoid as many undesirable influences as possible, the sequential presentation format was modified to, what we will name, a *parallel* presentation format.
In the parallel presentation format, a number of words are gated at the 'same' time, i.e. in parallel. Consequently, the gated stimuli belonging to a single word are not presented immediately after each other, but are interspered by speech fragments of other words. Given a sufficient large amount of words gated in parallel, as was the case in this experiment, it should be impossible for the listener to determine from the arrangement

Table 1. Duration in msec of the 3 vocalic segments for all words. CV = initial vocalic transition, V = steady-state part of a vowel, VC = final vocalic transition , total = the sum of CV, V and VC.

word	CV	V	VC	total	word	CV	V	VC	total
bal	10	137	24	171	bɑl	10	110	26	146
bas	11	158	31	200	bɑs	11	84	21	116
hak	21	132	22	175	hɑk	20	55	23	98
hal	22	102	24	148	hɑl	22	58	24	104
kap	8	150	15	173	kɑp	9	67	25	101
kas	8	156	26	190	kɑs	8	78	17	103
mat	10	139	21	170	mɑt	11	61	41	113
pal	8	143	26	177	pɑl	5	92	25	122
pap	14	139	36	189	pɑp	15	68	31	114
staf	13	159	17	189	stɑf	13	79	17	109
tak	11	143	22	176	tɑk	18	70	26	114
zak	25	135	42	202	zɑk	23	60	23	106
mean	13.4	141.1	25.5	180.0	mean	13.8	73.5	24.9	112.2

word	CV	V	VC	total	word	CV	V	VC	total
bom	9	126	25	160	bɔm	10	78	25	113
bon	11	157	30	198	bɔn	12	92	22	126
bot	13	121	27	161	bɔt	11	57	32	100
dop	9	116	31	156	dɔp	9	65	23	97
kok	8	117	16	141	kɔk	9	55	14	78
kop	10	123	22	155	kɔp	13	54	19	86
mot	11	120	28	159	mɔt	11	70	28	109
pol	11	162	27	200	pɔl	16	106	26	148
pot	8	125	26	159	pɔt	9	69	22	100
rok	11	102	16	129	rɔk	13	58	31	102
ton	8	104	26	138	tɔn	8	58	23	89
zon	21	142	25	188	zɔn	21	91	25	137
mean	10.8	126.3	24.9	162.0	mean	11.8	71.1	24.2	107.1

word	CV	V	VC	total	word	CV	V	VC	total
χest	8	122	18	148	χIst	19	66	17	102
kep	8	110	18	136	kIp	13	45	19	77
knep	10	108	28	146	knIp	10	52	23	85
krek	21	113	23	157	krIk	20	68	29	117
mes	11	112	19	142	mIs	11	45	23	79
pen	14	122	24	160	pIn	14	54	24	92
prek	10	109	22	141	prIk	9	43	17	69
rep	10	114	33	157	rIp	10	67	29	106
spen	11	115	24	150	spIn	9	45	24	78
strep	29	113	26	168	strIp	29	59	27	115
tek	8	106	31	145	tIk	15	46	29	90
ten	8	134	25	167	tIn	7	66	24	97
mean	12.3	114.8	24.3	151.4	mean	13.8	54.7	23.8	92.3

word	CV	V	VC	total	word	CV	V	VC	total
bϕk	14	115	30	159	bœks	11	56	23	90
bϕl	8	134	24	166	bœl	6	110	23	139
kϕs	9	134	23	166	kœs	9	78	18	105
lϕs	10	127	18	155	lœs	23	71	23	117
pϕk	9	109	18	136	pœk	10	57	18	85
pϕl	14	124	26	164	pœl	12	83	29	124
pϕt	12	122	22	156	pœt	15	64	19	98
rϕk	10	125	26	161	rœk	10	57	31	98
rϕs	10	167	22	199	rœs	11	77	23	111
slϕr	10	110	24	144	slœrf	9	87	24	120
tϕχ	15	114	21	150	tœχt	18	43	20	81
zϕχ	20	129	15	164	zœχt	20	55	15	90
mean	11.8	125.8	22.4	160.0	mean	12.8	69.8	22.7	104.8

of the stimuli itself, whether two acoustic-phonetic similar word fragments stem from one or from two different words.
The actual presentation sequences were prepared in the following way. The experimental words were divided into six different subsets. Each subset contained half of the experimental words: of two vowel opposition pairs all the words containing the long vowels as well as those containing the short vowels of the two remaining oppositon pairs. Thereby, six different combinations are possible, each of which was realised in a subset.
A different presentation sequence was prepared for each of the six subsets. It was accomplished in the following manner. The 48 words of a subset were randomly divided into 12 groups of 4 words. The stimulus sets of these words were then entered into the presentation sequence in 12 steps. At the first step, the stimulus sets of the first group of words were placed into the sequence. The first gates of these words were taken down. At the next step the second group of 4 words was added. Thus, 8 words had been gated in parallel by that time. The order of the appropriate stimuli (i.e. the second gates of the first group of words and the first gates of the second group of words) was randomized once more and taken note of. This procedure was repeated until all 12 groups had been introduced into the sequence.
Similarly, a partial presentation sequence was made for the filler items. Both the sequence of experimental stimuli and the one of filler stimuli were combined in such a way that at most two experimental stimuli followed each other. Each of the six combined sequences were preceded by the first 15 filler stimuli. These 15 stimuli served as practice items. Each of the resultant presentation sequences contained approximately 815 stimuli.
A presentation sequence is characterised by a continuously varying number of words being gated in parallel. At the beginning of the sequence the number of words increased steadily step-by-step. Once the complete set of gated stimuli of a word had been presented, these were then withdrawn from the sequence, thus decreasing the number of words being gated in parallel.

Stimulus tapes

Twelve audio tapes were prepared from the digitally stored stimuli. The stimuli were low pass filtered at 4.5 KHz and made analog by using a 12 bit D/A convertor, a Revox A77 2-track tape recorder and Agfa PEM268 tapes at a recording speed of 3.75 inch/sec. Six of the tapes contained the stimuli in the Silence (S) condition, while the remaining six tapes contained the stimuli in the Noise-Replaced (NR) condition.

Procedure

The stimulus tapes were presented binaurally via a Revox A77 2-track tape deck, a Sansui AU-D22 amplifier and Senheisser HD424 headphones at a comforting listening level. Each stimulus tape was heard by three listeners. Subjects were tested in groups of up to three in a quiet room.
Subjects were told that they would hear one word at a time and that, occasionally, a considerable portion of the word had been cut off (in the S-condition) or had been made unintelligible by noise (in the NR-condition). The subjects were instructed to write down the word they had heard after each presentation of a stimulus. They were required to give a response, however unsure they might feel about the stimulus word. The tape was played uninterruptedly. The duration of the interstimulus interval was four and a half seconds.
Each page of the response sheet consisted of 70 numbered lines, one for each response. One cue tone appeared after every tenth stimulus for orientation purposes, whereas two

cue tones after every seventeenth stimulus signalled the listeners to turn to the next page of their response sheet.
A short break was allowed after the presentation of 280 stimuli and another one after a further 280 stimuli. An entire listening session lasted about one hour and a half.

Subjects

Subject were 36 students, who were paid for their services. All subjects were native speakers of Dutch, with no reported hearing loss.

RESULTS AND DISCUSSION

Preliminary remarks on data analysis

The data of each vowel opposition pair have been analyzed separately. This approach was chosen for because of several reasons. Firstly, the immediate consonantal context was only controlled for within the set of words of each vowel opposition, but not among the four vowel oppostions. Secondly, the frequency of occurrence of the different vowels varied considerably (van den Broecke, Aerts, Reizevoort, Veenhof, Lammens, & Elstrot, 1987). Thirdly, the frequency of the words was not controlled for. Fourthly, only the speech materials of a single speaker were used. Any perceptual differences among vowel oppositions in this experiment might, therefore, be attribut - able to either a difference in the actual acoustical realization of the vowels by the speaker, or to any other difference among the vowels. A detailed comparison of the four vowel oppositions is therefore probably not meaningful.

Overall effects

The question of to what extent vowel length and type of gated stimulus have an effect can be answered by examining the overall effects of timing on vowel identification. The gating task commonly provides two global measures for this, i.e. the isolation point and the recognition point. We could not calculate recognition points, as we did not ask for confidence ratings in this experiment. The reason for this is easy to understand. The listeners were instructed to respond to each presented stimulus with a word. Confidence ratings would then reflect how sure a listener is about the word. However, as we are interested in vowel identification, confidence ratings reflecting word recognition points are of no use to us.
A vowel of a word was said to be isolated at the gate at which a subject correctly guessed the identity of a vowel and did not subsequently change his or her mind. The mean vowel isolation point for a subject was taken to be the mean amount of vowel signal of the ten (out of twelve) earliest isolated vowels. The analyses of the vowel isolation points are displayed in Figure 2. All vowel isolation points are given relative to the beginning of the vowel. Two effects are shown in Figure 2. The first one is the difference between the long and short vowel (Type of Vowel) of each opposition pair. The second one is the effect of the different Type of Gate (S- versus NR-condition).
Separate analyses of variance (ANOVAs) have been performed on the subject mean isolation points for each vowel opposition pair. All F-values reported are significant at the .05 level or above, unless otherwise noted.

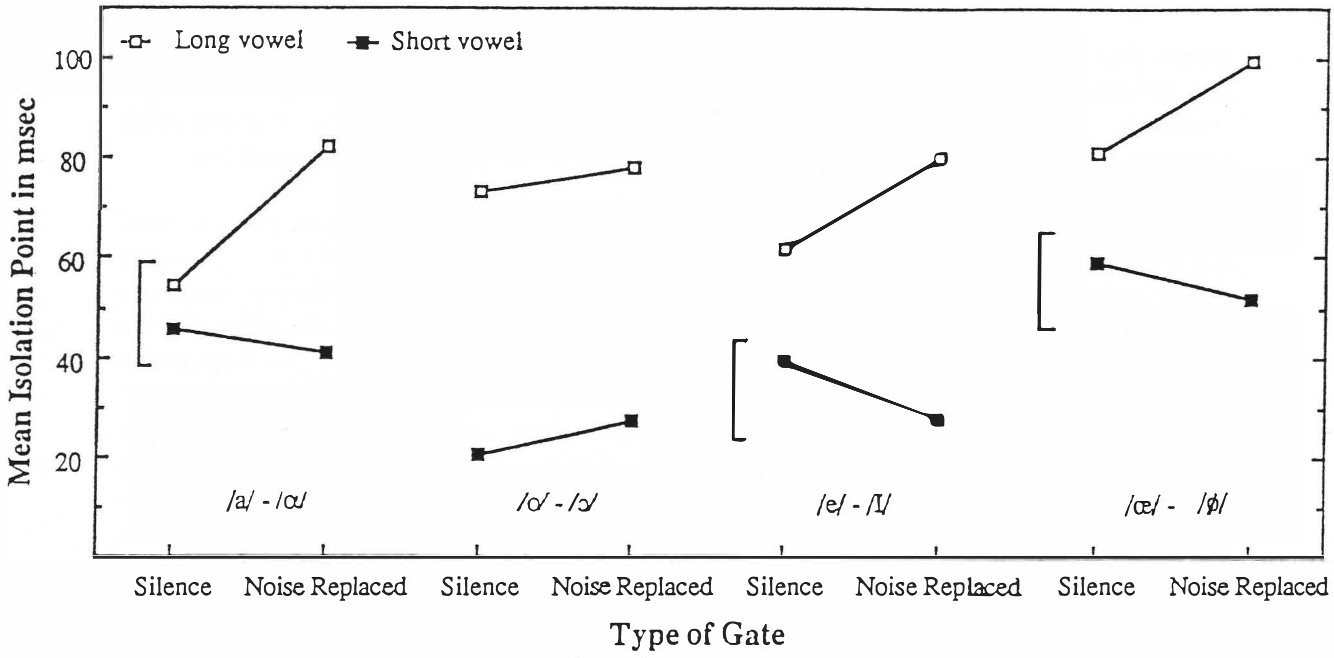

Figure 2. Isolation points computed for each of the vowels of the 4 vowel opposition pairs for both types of gated stimuli (Silence and Noise Replaced). Open squares indicate lax vowels; closed squares indicate tense vowels. Left square brackets enclose means that do not differ significantly by a Newman-Keuls test with $p < .05$.

Vowel opposition pair /a/ - /ɑ/ Firstly, a main effect of Type of Vowel (long versus short) was found [F(1,32) = 21.4] with the isolation point of the long vowel /a/ lagging behind by 24 msec averagely on the isolation point of the short vowel /ɑ/. There was also a main effect of Type of Gate [F(1,32) = 4.32]. The isolation point of the stimuli in the S-condition fell 11 msec earlier than in the NR-condition. Both factors interacted significantly [F(1,32) = 8.73]. Post hoc comparisons (Newman-Keuls) indicated that the mean isolation point of the long vowel /a/ in the NR-condition differed significantly from the three others. See Figure 2. The significant main effect of Type of Vowel was mainly caused by the difference between the isolation points of the short and the long vowel in the NR-condition. This calls into question the importance of vowel duration for the isolation of the long vowel /a/. If we assume that there is a general masking effect of noise, it is hard to explain why the noise merely affected the isolation point of the long vowel.

Vowel opposition pair /o/ - /ɔ/ A different pattern of results emerged from the analysis of the vowel pair /o/ - /ɔ/. See Figure 2. The results indicated a strong main effect of Type of Vowel [F (1,32) = 178]. The difference between the isolation points of the long and short vowels averaged about 52 msec. The main effect of Type of Gate was not significant [F (1,32) = 2.26, n.s.], nor was the interaction between both factors [$F < 1$]. Vowel length is thus a cue to discriminate between the long and short vowel and moreover equally important for both types of gated stimuli. It is remarkable that the noise of the NR stimuli did not display a negative effect, as this clearly was the case with the long vowel /a/.

Vowel opposition pair /e/ - /I/ The factor Type of Vowel showed a significant main effect [F(1,32) = 64.9] with the average isolation point of the short vowel /I/ occurring some 38 msec earlier than the average isolation point of the long vowel /e/. There was no significant main effect of the factor Type of Gate [$F < 1$], but the interaction of both factors was [F(1,32) = 10.1] The isolation points of the vowels /I/ and /e/ differed about 23 msec in the S-condition and enlarged to 53 msec in the NR-condition. See Figure 2. The length of the vowel segment is thus a cue to distinguish between /I/ and /e/ and, moreover, a cue for the vowel stimuli in the NR-condition.

Vowel opposition pair /ϕ/ - /œ/ The results of the analysis of the vowel pair /ϕ/ - /œ/ are comparable to those of the vowel opposition pair /e/ - /I/. See Figure 2. A main effect of Type of Vowel was found [F(1,32) = 32.7]. The average difference between the isolation point of the vowels /ϕ/ and /œ/ amounted to 35 msec. The effect of the second factor (Type of Gate) was not significant [F < 1]. Finally, the interaction between the factors Type of Vowel and Type of Gate was significant [F(1,32) = 4.39]. The difference between the mean isolation points of the long and short vowel was far smaller in the S-condition (21 msec) than in the NR-condition (48 msec). As there was no significant difference between the isolation point of the short vowel /œ/ in both Type of Gate conditions, it is clear that the noise of the stimuli in the NR-condition hampered early isolation of the long vowel /ϕ/.

It was pointed out earlier that a detailed comparison of the four vowel opposition pairs is not meaningful. Accordingly, only a global comparison of the four opposition pairs will be outlined.
The analyses decribed above give evidence for the view that the difference in duration of the long and short vowels is a cue to distinguish between the two vowels of an opposition pair. The vowel pair /a/ - /α/ in the S- condition takes the only exception to this. This vowel opposition pair has been studied extensively and all the experiments reported in the literature show a definite effect of vowel length on the identification of the vowels /a/ and /α/. See for example Nooteboom and Doodeman (1980).
The influence of noise in the NR-stimuli is contradictory. On the one hand, the noise delayed the early isolation of the long vowels /a/, /e/ and /ϕ/, but had no effect on the isolation point of the vowel /o/. On the other hand, the noise did not affect the isolation points of any of the four short vowels. These data indicate that the noise did <u>not</u> mask the spectral contents of the speech fragments, but only influenced the perception of the duration of the vowel segment. This hypothesis is tested in the next section.

Effects of segment size of the long vowels on vowel identification

The overall effects indicated that vowel length is a cue to distinguish the difference between long and short vowels. Noise in the NR stimuli had a negative effect on the early isolation of the long vowels, yet had no effect on the short vowels.We suggested in previous section that the noise might have influenced the perception of the duration of the vowel segment. Given the observed overall effects, it could be assumed that the noise in the NR-stimuli enlarged the proportion of short vowel responses. This hypothesis can be evaluated by examining the number of long and short vowel responses offered as candidates to the long vowel stimuli at each gate. A combination of two effects should be ascertained to substantiate the hypothesis. Firstly, a larger number of long vowel responses to long vowel stimuli should be encountered in S-condition than in the NR-condition. Secondly, the reverse effect should be observed for the short vowel responses to the long vowel stimuli. If the first effect is significant and the second one is not, then we have no choice but to conclude that the noise masked the spectral contents of the speech fragments of the stimuli in the NR-condition.

Figures 3 to 6 show the proportion of long and short vowel responses given at each of the first seven gates to, respectively, the long vowels /a/, /o/, /e/ and /ϕ/ in the combination of Vowel-segment-size with both Type of Gate conditions. These figures yield a detailed picture as to what extent the size of the vowel segment had an effect upon the choice of either the long or the short member of a vowel opposition pair. With the exception of the vowel /o/, it is evident that the proportion of long and short vowel responses under the S- and NR-conditions is as hypothesized. This has been confirmed statistically. For each of the four long vowels, two separate ANOVAs were performed: One analysis on the number of correct, long responses and one on the number of incorrect, short responses. The number of these responses at the gates 2 to 6 (gates 3 to 7 of the vowel /ɸ/) given by each subject were calculated. These values were entered into ANOVAs with the two Type of Gate conditions crossed with Vowel-segment-size. Subjects were considered as a random effect. Figure 3 clearly demonstrates a lack of difference between the S- and NR-condition for the vowel /o/. Therefore, only the the data of the three remaining long vowels are reported here. The results are as follows:

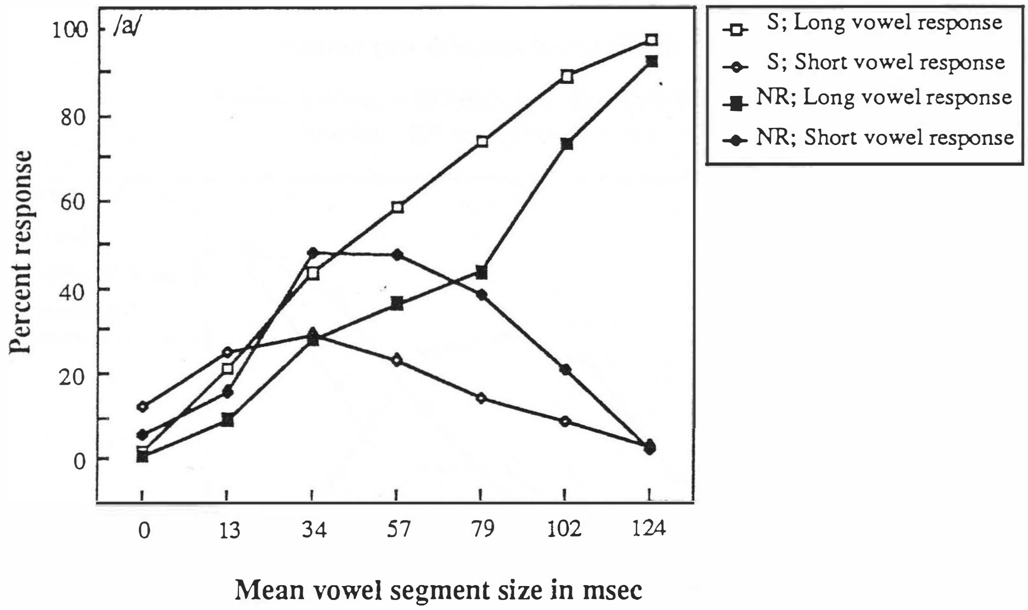

Figure 3. Percent /a/ and /ɑ/ responses by gate to /a/ stimuli for the stimuli in the S and NR condition

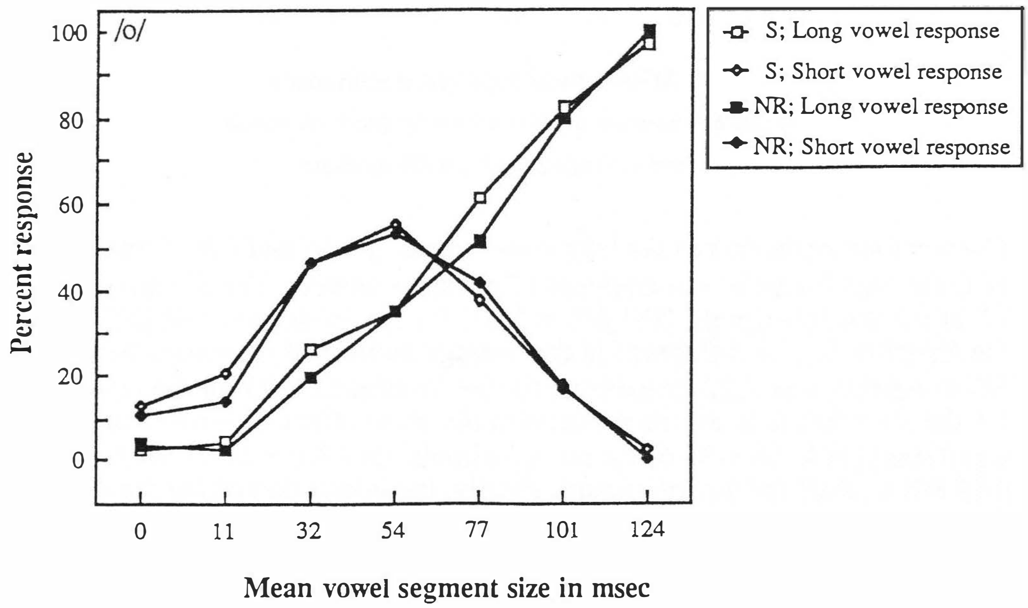

Figure 4. Percent /o/ and /ɔ/ responses by gate to /o/ stimuli for the stimuli in the S and NR condition

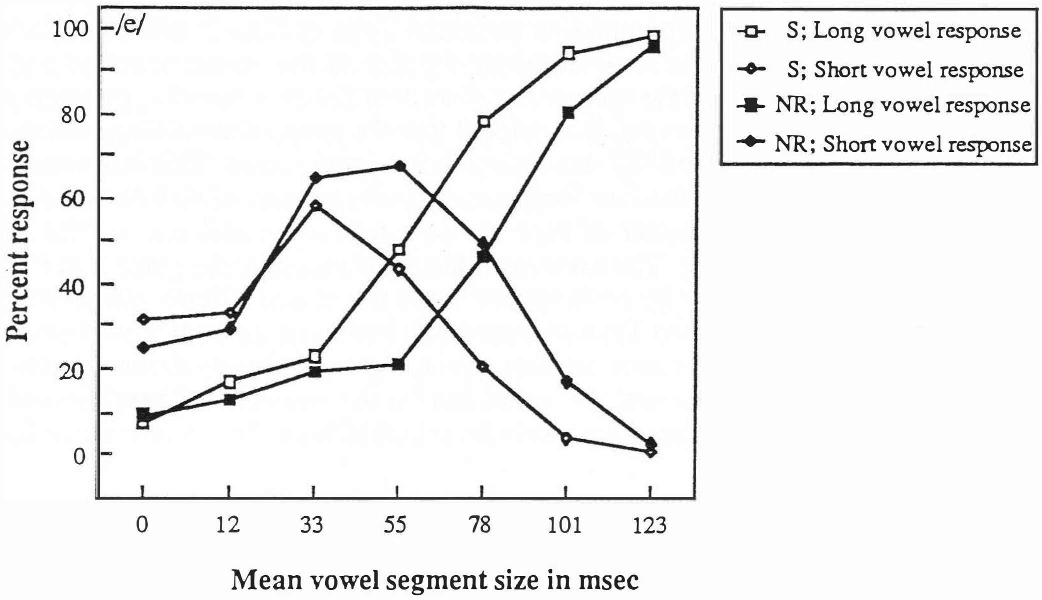

Figure 5. Percent /e/ and /I+ɛ/ responses by gate to /e/ stimuli
for the stimuli in the S and NR condition

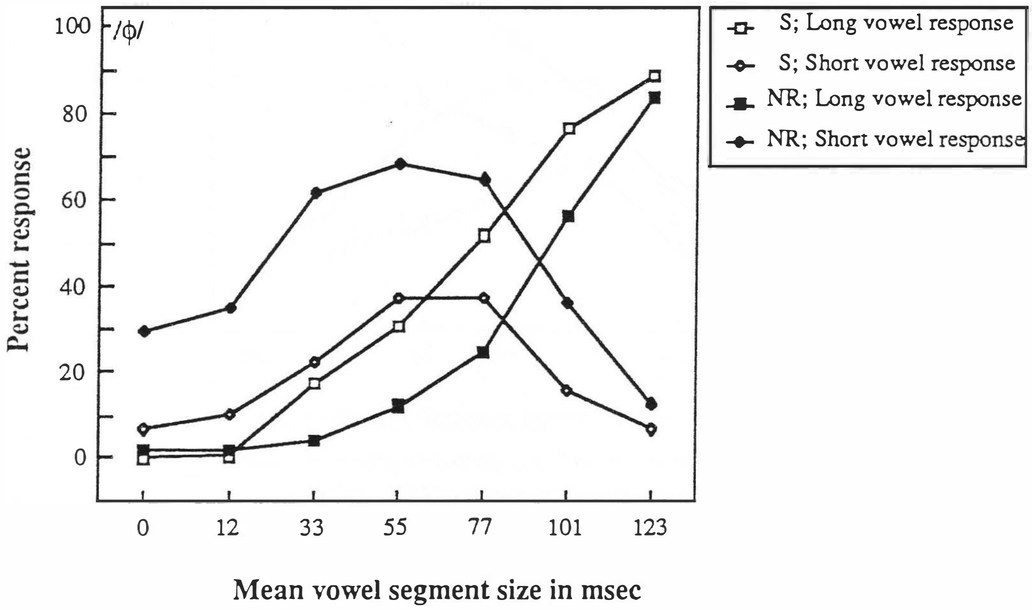

Figure 6. Percent /ϕ/ and /ʏ/ responses by gate to /ϕ/ stimuli
for the stimuli in the S and NR condition

Correct long responses to the long vowel stimuli /a/, /e/ and / ϕ/ A main effect of Type of Gate was found in the analyses of all three vowels. The F-values are: [F(1,80) = 23.2] for the /a/-stimuli, [F(1,80) = 22.0] for the /e/-stimuli and [F(1,80) = 8.38] for the /ϕ/-stimuli. The difference in the average number of responses between the S- and NR-condition was 2.22 (responses) for the /a/-stimuli, 1.93 for the /e/-stimuli and 1.53 for the /ϕ/-stimuli.In all three analyses the main effect of Vowel-segment–size was significant [F(4,80) = 34.6] for the /a/-stimuli, [F(4,80) = 61.2] for the /e/-stimuli and [F(4,80) = 36.0] for the /ϕ/-stimuli. Finally, the interaction of the factors Type of Gate and Vowel-segment-size was significant for the /e/-stimuli [F(4,80) = 2.99], yet unsignificant for the /a/-stimuli [F < 1] and the /ϕ/-stimuli [F(4,80) = 1.12, n.s.].

Incorrect short responses to the long vowel stimuli /a/, /e/ and /ϕ/ The analyses showed that for all three vowels, the average number of short responses was significantly larger in the NR-condition than in the S-condition (/a/-stimuli: [$F(1,80) = 10.3$], /e/-stimuli: [$F(1,80) = 23.2$] and the /ϕ/-stimuli: [$F(1,80) = 43.0$]). The difference in average number of responses to the the vowels /a/, /e/ and /ϕ/ was, respectively, 1.65, 2.05 and 3.00 (responses). As could be expected, the main effect of Vowel-segment-size was significant in all three analyses (/a/-stimuli: [$F(4,80) = 5.29$], /e/-stimuli [$F(4,80) = 28.2$] and the /ϕ/-stimuli: [$F(4,80) = 18.7$]). The interaction between the two factors was not significant in either analysis (/a/-stimuli: [$F(4,80) = 2.39$, n.s.], /e/-stimuli [$F(4,80) = 2.14$, n.s.] and the /ϕ/-stimuli: [$F(4,80) = 2.30$, n.s.]).

The results of the analyses support the above mentioned hypothesis. In comparison with the S-condition, the presentation of long vowel segments of the stimuli in the NR-condition led to a change of the ratio of long and short vowel responses in the hypothesized way. Due to the fact that the two members of a vowel opposition pair resemble each other spectrally, we conclude that the noise in the NR-stimuli did not mask the nontemporal, spectral aspects of the vowels. Yet, both types of gated stimuli differently affected the perceived duration of the long vowels. Nevertheless, we think that it is incorrect to conclude that the noise masked the duration of the vowel fragments. Duration (of vowel segments) is very likely to be perceived in a relative and not in an absolute manner. Therefore, it is just as possible that the listeners perceptually overestimated the duration of the long vowel segments in the S-condition, as underestimated it in the NR-condition. The present experiment offers no means to discriminate between these two alternatives.

CONCLUSION

In this study, we have tried to answer a specific empirical question: Is the gating paradigm, developed by Grosjean (1980), a suitable method for studying long-short vowel opposition pairs. The results affirm this. It was shown that the isolation points of the long vowels lagged behind those of the short vowels and that the presentation of short segments of long vowels led to confusions of long vowels with their short counterparts.

We compared the standard method of generating gated stimuli to one in which portions of words, that are deleted in the standard method, were replaced by envelope-shaped speech noise. The isolation points of the long vowels of these stimuli occurred later than those of the standard generated stimuli. The isolation points of the short vowels of the two types op gated stimuli, however, did not differ. After analyzing the number of long and short responses given as candidates to the shorter segments of the long vowels, we concluded that the two types of stimuli affected the perceived length of the long vowel segments in a different manner. As we have no objective criterion for a 'real' perceptual length of a vowel segment, we can not conclude that the noise masked the speech fragments of these stimuli.

We believe that this extension of the gating technique is a useful one. It allows one to gate several words in a single utterance (as was done by Salasoo and Pisoni,1985) and it opens the possibility to embed the gated word in an utterance.

ACKNOWLEDGMENT

This research is supported by Grant 300-161-32 from the Foundation for Linguistic research to Louis C.W. Pols. The Foundation of Linguistic research is funded by the Netherlands organization for the advancement of pure research (Z.W.O.).

REFERENCES

Bergem, D.R. van, & Drullman, R. (1985) The influence of contextual speech rate on the identification of Dutch vowels in normally and rapidly spoken tVt utterances. IFA-report nr. 77 (in Dutch)

Broecke, M.P.R. van den, Aerts, A., Reizevoort, J., Veenhof, T., Lammens, J., & Elstrot, M. (1987) Type- and token-frequencies of wordclasses, phonemes and phoneme pairs in Dutch. PRIPRU, 12, 1-15.

Cotton, S., & Grosjean, F. (1984) The gating paradigm: A comparison of succesive and individual presentation formats. Perception and Psychophysics, 35, 41-48.

Grosjean, F. (1980) Spoken word recognition and the gating paradigm. Perception and Psychophysics, 28, 267-283.

Grosjean, F. (1985) The recognition of words after their acoustic offset: Evidence and implications. Perception and Psychophysics, 38, 299-310.

Johnson, T.L., & Strange, W. (1982) Perceptual constancy of vowels in rapid speech. J. Acoust. Soc. Am., 72, 1761-1770.

Nakatani, L.H., & Dukes, K.D. (1973). A sensitive test of speech communication quality. J. Acoust. Soc. Am., 53, 1083-1091.

Nooteboom, S.G., & Doodeman, G.J.N. (1980) Production and perception of vowel length in spoken sentences. J. Acoust. Soc. Am., 67, 276-287.

Salsoo, A., & Pisoni, D.B. (1985) Interaction of knowledge sources in spoken word identification. Journal of Memory and Language, 24, 210-231.

Tyler, L.K. (1984) The structure of the initial cohort: Evidence from gating. Perception and Psychophysics, 36, 417-427.

Tyler, L.K., & Wessels, J. (1983) Quantifying contextual contributions to word recognition processes. Perception and Psychophysics, 34, 409-420.

Tyler, L.K., & Wessels, J. (1985) Is gating an on-line task? Evidence from naming latency data. Perception and Psychophysics, 34, 217-222.

Warren, P., & Marslen-Wilson, W. (1987) Continuous uptake of acoustic cues in spoken word recognition. Perception and Psychophysics, 41, 262-275.