Initial /b/, /p/, /d/, /t/ - A Listening Test ·)

J.W. Meijers

·) The exploratory investigation reported in this paper was carried out in partial fulfilment of the requirements for a Master's Degree in General Linguistics and Phonetic Sciences at the University of Amsterdam.
I wish to thank Mrs. H. Deighton of the Institute of Phonetic Sciences for the many stimulating discussions on the subject of this paper.

Contents

Summary

Dutch listeners were asked to judge whether the initial plosives /b̥/ and /d̥/ belonged to /b/, /d/ or to /p/, /t/ respectively.
The notation b̥ and d̥ is used to indicate that
1. the vocal cords do not vibrate before the release of the plosive;
2. the sound is non-tense, or lax (cf. esp. Kim, 1965).
In spite of the fact that /b/ and /d/ are non-tense, the listeners class the sounds with the /p/ and /t/ phonemes. The absence of voicing seems sufficient for them not to choose for /b/ and /d/.
It is suggested that the voice feature is the primary cue for Dutch listeners in the perception of initial plosives.

1 Introduction

1.1 Voiced-Voiceless and Lax-Tense

The terms voiced and voiceless are used in the following sense:
A plosive is called voiced when the vocal cords vibrate during (part of) the closure, causing a buzz which is visible on the oscilloscope as regular oscillation.
A plosive is called voiceless when no vibration of the vocal cords takes place before the release.
As, for my purposes, a term is needed to cover phenomena outside the voiced-voiceless dimension that may play a part in production and perception, I feel justified in using the terms lax and tense for initial plosives, basing myself on measurements done by Kim (1965) in this field. The terms are thus used negatively as a working hypothesis. It seems undesirable to introduce yet another set of terms. ') Further research into the exact nature and role of these factors, if any, in the production and perception of initial plosives remains necessary.
The notation /b/ and /d/ serves to indicate that the plosive is voiceless but non-tense.
Writers are by no means in agreement on the relative importance they attach to the components in the production of (plosive) consonants. Stetson (1951), for instance, maintains that "difference in pressure, expressed by the terms fortis and lenis, is more fundamental than the voicing of the consonants, and persists after the voicing distinction is lost". Jakobson and Halle (1956) also seem to regard the tensity distinction as the primary cue.
Lisker and Abramson (1964), however, regard the voiced-voiceless distinction as primary, and attempt to classify the

') Some writers use the terms fortis-lenis or strong-weak instead of tense-lax. In quotations the original words are kept.

plosives of 11 languages in terms of the duration of voicing. Kim (1965) argues that voicing and tensity are autonomous cross-cutting features of plosives, and Fant (1967) suggests that the voiced-voiceless distinction would be just as useful as the lax-tense distinction. Moulton (1962) distinguishes the two Dutch plosive series merely by calling them fortis and lenis.

The subject of this study is the role played by the feature voice in the perception of initial English and Dutch plosives, when judged by Dutch listeners.
In the phonemic system of British English there exist the oppositions /b/ - /p/, /d/ - /t/, and /g/ - /k/. Dutch only has /b/ - /p/ and /d/ - /t/.
Unable to investigate personally the validity of the tense-lax dimension, I shall follow the general description of voiced plosives being lax and voiceless plosives tense in most languages of the world, including English and Dutch. For Dutch plosives, for instance, Cohen c.s. (1969;81) state: "/p,t,k/ are tenser and have greater penetration than /b,d,g/". ')
The data of the present investigation confirm the observation that in English initial /b/ and /d/, voice may not set in until the end of the closure or even be completely absent during the closure. Among the English items with initial /b/ and /d/ occurred a fair number in which vocal cord vibration begins after the release. This is never the case with the three Dutch speakers. The mean values of voice onset time for initial Dutch /b/ and /d/ in my data are -88 and -92 Msec. respectively. The three informants varied little in this respect. Lisker and Abramson (1964) found -85 and -80 Msec. with one native speaker of Dutch.
In the case of the English items with vocal cord vibration setting in after the release, the plosive is nevertheless lenis (Gimson,1962;147), so that the following sets of plosives occur in the words spoken by informants for this study: voiced /b/, /d/, devoiced /b̥/, /d̥/, and voiceless /p/, /t/, with features as shown in Fig. 1.

') The original runs as follows: /p,t,k/ zijn meer gespannen en hebben een groter doordringingsvermogen dan /b,d,g/.

1.2 Listeners divided into two groups

It was thought possible that perception of English plosives might be influenced by the Dutch listener's familiarity with the English language.
The listening test was therefore carried out in two groups, the one group having had no previous knowledge of English, the other consisting of students reading English at university level.

2 Method

2.1 Subjects

2.1.1 Speakers

The voices of six male speakers were recorded. Three of the informants were native speakers of English, having a fairly uniform Southern British English pronunciation, without dialectal variants.
The three others were Dutchmen with again a reasonably non-dialectal pronunciation of Dutch.
The speakers were unaware of the purpose of the test.

2.1.2 Listeners

There were 130 listeners in all, of both sexes. The following subdivision was made: the larger group of 70 subjects consisted of primary school pupils who had not yet been taught English. Care was taken not to include any who had become acquainted with English in some other way. The ages varied between 10 and 13, with an average of 11.5. The school was situated in a small town near Amsterdam, with pupils of different backgrounds. The second group consisted of 60 students reading English at university level. The greater part were in their early twenties.

2.2 Procedure

2.2.1 Recording

The recordings were made with professional apparatus at the Institute of Phonetic Sciences in the University of Amsterdam. Low print Scotch tapes were used, speed 7.5 in/sec.
The informant was seated in a sound-treated booth and asked to read aloud the word shown to him on a card. In order to approach a normal enunciation as closely as possible and also

to allow for adjustments in recording-level to be made, the speaker first read a short prose-fragment before going on to the cards.
The six informants each read 16 words in a random order, different again with each speaker.

Table 1

Words spoken by English informants

bip	dip	pip	tip
boot	doot	put	toot
buck	duck	puck	tuck
balloon	dakota	paternal	tattoo

Table 2

Words spoken by Dutch informants

bip	dip	pip	tip
boet	doet	poet	toet
bak	dak	pak	tak
beleefd	de	petoet	te

The vowels following the initial plosives on the one hand approach the extremes of the vowel diagram and on the other hand have a near approximate in the other language: Table 3.

Table 3

Vowels in the words spoken

English	Dutch
/i/	/i/
/u/	/u/
/ʌ/	/ɑ/
/ə/	/ə/

2.2.2 Measuring of voice onset time

By means of an oscilloscope coupled to a segmentator, the initial plosives were measured on voice onset time. The oscilloscope was first checked with a tone of 1.000 cps, from a tone-oscillator.
Voice onset time (VOT) will be expressed in milliseconds with the moment of release taken as zero. A voice lag will thus be positive, a voice lead negative. (Cf. Figures 2, 3, and 4)
The 96 words were then reduced in length in such a way that items were left consisting of an initial plosive + vowel sound of about 100 Msec. In some cases of C + /ə/, the vowel sound, having been in unstressed position, is shorter.
The resulting 96 items were duplicated and randomised.
The final tape contained the following components:

1. 192 stimuli in a random sequence, each stimulus occurring twice;
2. 4 seconds of silence between successive stimuli;
3. an announcement of the number of each stimulus.

The playing time was 30 minutes.
Every effort was made to maintain a constant level of loudness. Inevitably, however, in the case of some of the C + /ə/ stimuli with their unfavourable signal-to-noise ratio, the result is less satisfactory.

/d/ in doet (Dutch)
VOT: -126 msec.
1 and 2 are first and second parts of the same sequence.

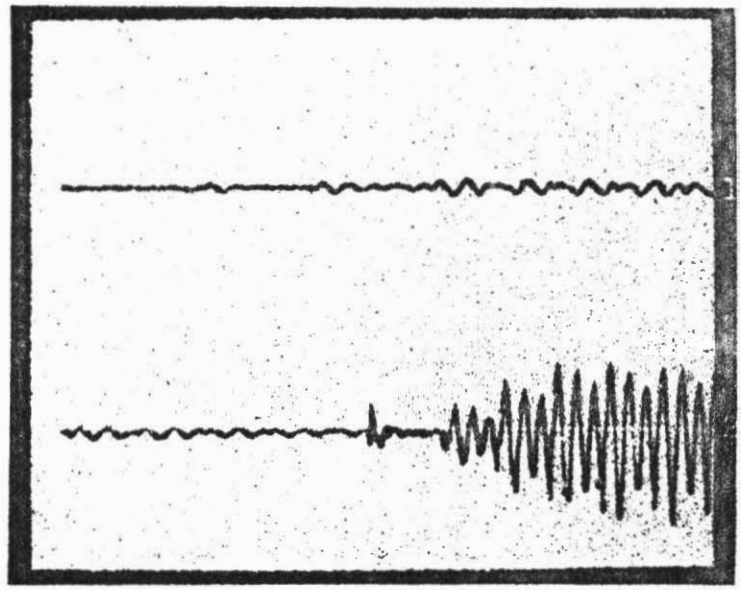

Fig. 2.

/d/ in doot (English)
VOT: 14 msec.

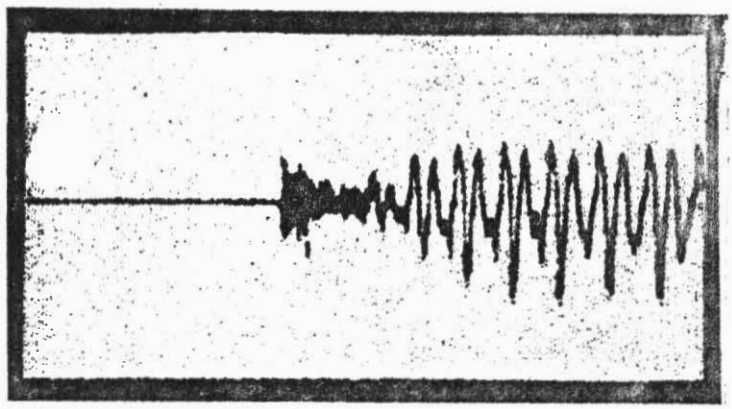

Fig. 3.

/t/ in toet (Dutch)
VOT: 18 msec.

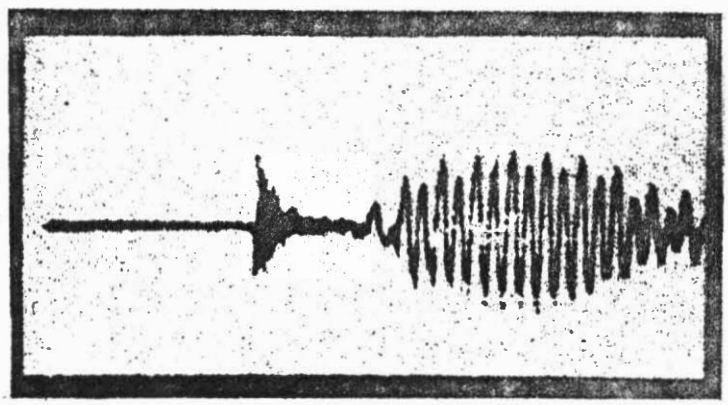

Fig. 4.

/t/ in toot (English)
VOT: 42 msec.

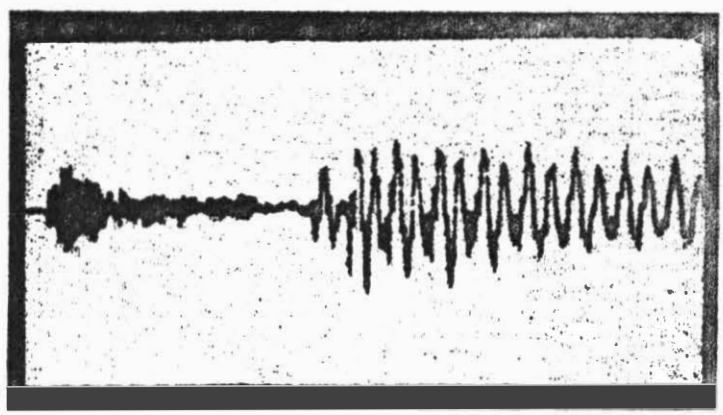

Fig. 5.

2.2.3 Scoring - how?

The subjects were asked to score on a specially prepared scoring-sheet. They were given a choice between two possibilities. (Cf. Table 4) In view of the fact that direct scoring might prove too difficult for the primary school pupils, it was decided to print the two possibilities and require the listeners to underline their judgment. This proved to be a very necessary and practicable arrangement for this particular test, for which we are only interested in the decision voiced/voiceless and not in the identification of the consonant proper.

Table 4

Section of the scoring sheet

1.	ti	di	51.	te	de
2.	pa	ba	52.	pi	bi
3.	pi	bi	53.	pu	bu
	etc.			etc.	

2.2.4 Scoring - where?

The listening test was carried out in ordinary reverberant rooms. A Tandberg taperecorder, model 14, was used to play the final tape. The recorder appeared to be very satisfactory in the rooms used.

A. The group of primary school pupils

The test took place in classrooms where the pupils ordinarily attended their lessons. It was considered an advantage for them to remain in familiar surroundings.
The tape was played for two groups of 35 pupils each, with

two breaks of about five minutes. The listeners first received instructions on what was expected from them.
The scoring presented no difficulties.

B. The group of students

The test took place in the manner described above, the only difference being that the groups consisted of not more than 10 listeners at a time.
It was found unnecessary to have breaks.

3 Results and Discussion

3.1 Difference in Scoring between the two groups

Table 5

Number and percentage of voiceless judgments for the two groups Group 1: primary school pupils. Group 2: students.

	Grand total (voiced & voiceless) per item	Voiceless scorings per item	
		Mean number	Mean percentage
Group 1	140	115.8	83 per cent.
Group 2	120	102.2	85
Groups 1 & 2	260	218.0	84

From Table 5 it appears that there is no significant difference in scoring. There is, in fact, a striking agreement. The distinction between the two groups will no longer be made in the following and observations will be based upon results of 130 subjects, who each scored on 192 items, yielding a total of 24,960 judgments. The full data of the computer output are available at the Institute of Phonetic Sciences.

3.2 Relation voice onset time - scoring

A distribution of voiceless judgments on a time basis leads to the distinction of two clearly separate groups, as shown in Fig. 6.
For values of VOT per speaker and per item cf. Appendix 1.

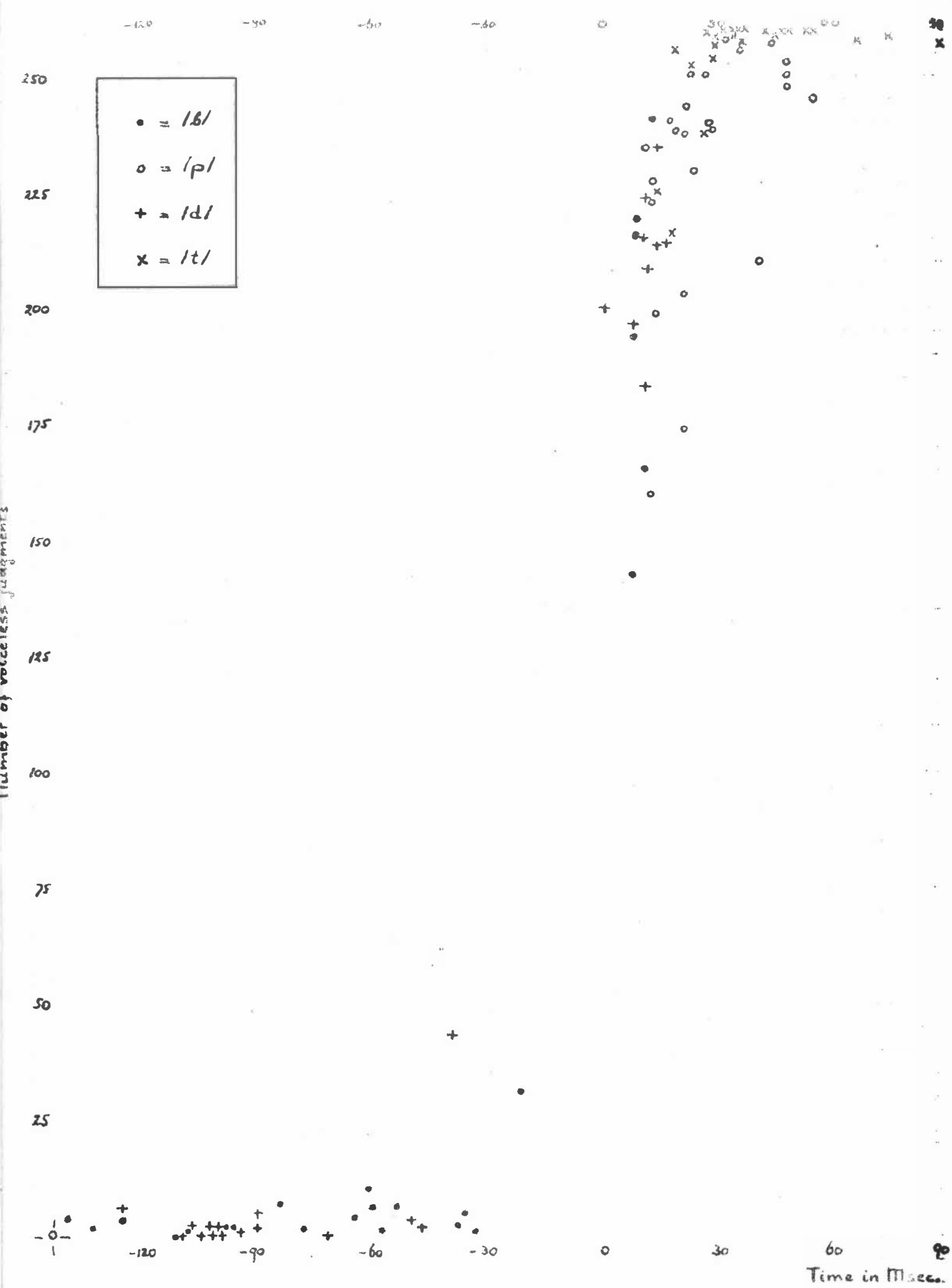

Fig. 6. Distribution of voiceless judgments on a time basis.

3.2.1 Items with voice onset time between -141 and -35 Msec.

Number of judgments: 32.
Voiceless scoring is negligible, ranging from 0 to 10, with an average of 2.8. This average is exclusive of the one exception, viz. item /dɑ/ of speaker 6 with a VOT of -41 Msec., which gave a comparatively high voiceless scoring of 43. Perception in this case, however, may have been influenced by the less satisfactory quality of the stimulus (Cf. 2.2.2).

3.2.2 Items with voice onset time between 0 and 90 Msec.

Number of judgments: 63.
Voiceless scoring is significant, ranging from 143 to 260, with an average of 234.
The data of the experiment give only one instance of an item with a VOT in the remaining range from -35 to 0 Msec., which is insufficient to establish the way in which the transition between the two groups takes place. It is nevertheless striking that perception of a voiced plosive or a voiceless one should be decided in the comparatively short range from -35 to 0 Msec.
These findings strongly corroborate the results of Slis and Cohen (1969,p.84), obtained in a commutation test: "voiced judgments correlate well with a voice lead and voiceless ones with a voice lag". Their figures show a 100 per cent voiced judgment on segments with (synthetic) voice lead variable from -200 to -40 Msec. and a decrease in voiced judgment after that. This, too, is borne out by the data in the present study.

3.2.3 Relation voice onset time - scoring within the two groups

No consistent relation was found between voice onset time and scoring within the two groups mentioned in 3.2.1 and 3.2.2. The scoring on some labials as shown in Table 6 may serve as an example.

Table 6

Scoring for some stimuli with (nearly) the same VOT

Stimulus	Speaker	VOT	Scoring (voiceless)
/pɑ/	1	13	199
/pɑ/	3	12	228
/pi/	1	12	161
/pi/	3	11	235
/b̥i/	4	10	166
/pu/	2	12	223
/b̥u/	4	12	242

In spite of a practically identical VOT, scoring differs considerably, even for identical stimuli of different speakers.

3.2.4 Scoring on /t/

Scoring on /t/ is highly significant, ranging from 216 to 260, with an average of 254.5.

3.3 Items of English speakers with voice onset time of 0 Msec. or more

Among the items spoken by the English informants were 14 items with a voice lag and 1 item with a VOT of 0 Msec. (Cf. Appendix 1). The average for voiceless scoring on these stimuli was 205, or 78.8 per cent. (Cf. Appendix 3).
There is no significant difference in the perception of /b̥/ or /d̥/ respectively, nor does the vowel sound seem to influence judgment on the preceding plosive.
Scorings per speaker did not differ from the whole.

Conclusion

/b̥/ and /d̥/ take up a position intermediate between /b/, /d/ and /p/, /t/. In being voiceless they correspond with /p/, /t/,

in being lax they correspond with /b/, /d/.
When forced to make a choice, as in our experiment, a significant number of Dutch listeners judge the absence of voice to be more important than the absence of tenseness, and consequently class /b̥/, /d̥/ with /p/, /t/.
On this basis it would seem that in the perception of Dutch listeners the voiced-voiceless dimension is a primary cue in initial plosives, and the lax-tense dimension of secondary importance.

It should be noted that it is in no way intended to express an opinion on the relation English - Dutch.
The English plosives /b/, /d/ with voice lag were merely found useful in that the features voice and force of articulation ("tenseness") are not linked in the way found in Dutch initial plosives.
It is only the perception of initial plosives generally by Dutch listeners with which we have been concerned.

A sequel to the present study might be one in which values of voice onset time were compared for a larger number of initial English plosives. One of the three English informants consistently articulated /b/, /d/ with voice lead. Lisker and Abramson (1964) found the same with one of their four speakers. It might be interesting to trace this further.
Another study could occupy itself with the way in which the stimuli of the present investigation would be perceived by British listeners. It would seem that their perception of /b̥/ and /d̥/ might be different from the Dutch listener's, as the sounds are part of the phonemic system of English but not of Dutch. Of interest would also be their reaction to Dutch /p/ and /t/, particularly to those items which have the same VOT as English /b/ and /d/ (Fig.6).

Appendix 1

Values of voice onset time for the plosives contained in the listening test.

Speaker		/ɑ/	/i/	/u/	/ə/
1 (Dutch)	/b/	-110	-141	-113	-100 Msec.
	/p/	13	12	21	21
	/d/	-106	- 91	-126	-112
	/t/	22	26	27	28
2 (Dutch)	/b/	- 59	- 79	- 39	- 61
	/p/	16	18	12	20
	/d/	- 73	- 48	- 52	- 99
	/t/	14	28	18	33
3 (Dutch)	/b/	-134	-126	- 37	- 55
	/p/	12	11	31	27
	/d/	- 98	- 91	-113	- 99
	/t/	22	29	31	31
		/ʌ/	/i/	/u/	/ə/
4 (Engl.)	/b/	8	10	12	7
	/p/	54	47	47	21
	/d/	13	7	14	16
	/t/	74	90	54	35
5 (Engl.)	/b/	8	7	- 66	- 23
	/p/	22	26	26	23
	/d/	9	11	10	10
	/t/	53	38	42	33
6 (Engl.)	/b/	-100	- 86	- 35	- 63
	/p/	43	47	35	40
	/d/	0	-108	- 95	- 41
	/t/	66	48	47	45

Appendix 2

Average values of voice onset time for the plosives contained in the listening test.

The first line per speaker gives the average VOT, the second line states the two extremes found.

Speaker	/b/	/d/	/p/	/t/
1 (Dutch)	-116	-109	17	26
	-141/-100	-126/-91	12/21	22/28
2 (Dutch)	-60	-68	17	23
	-79/-39	-99/-48	12/20	14/33
3 (Dutch)	-88	-100	20	28
	-134/-37	-113/-91	11/31	22/31
4 (Engl.)	9	13	42	63
	7/12	7/16	21/54	35/90
5 (Engl.)	-19	10	24	42
	-66/7	9/11	22/26	33/53
6 (Engl.)	-71	-61	41	52
	-100/-35	-108/0	35/47	45/66

Appendix 3

Number and percentage of voiceless judgments for 15 items of English speakers.

Speaker	Plosive	Vowel	VOT	Voiceless judgments Number	Percentage
4	/b̥/	/ʌ/	8	217	83 per cent.
	/b̥/	/i/	10	166	64
	/b̥/	/u/	12	242	93
	/b̥/	/ə/	7	143	55
5	/b̥/	/ʌ/	8	220	85
	/b̥/	/i/	7	194	75
4	/d̥/	/ʌ/	13	214	82
	/d̥/	/i/	7	197	76
	/d̥/	/u/	14	235	90
	/d̥/	/ə/	16	215	83
5	/d̥/	/ʌ/	9	216	83
	/d̥/	/i/	11	209	80
	/d̥/	/u/	10	183	70
	/d̥/	/ə/	10	224	86
6	/d̥/	/ʌ/	0	200	77
	Grand Total			3075	78.8 per cent.

References

Cohen, A., Ebeling, C.L., Fokkema, K. en van Holk, A.G.F. (1969) Fonologie van het Nederlands en het Fries, 2nd ed., The Hague, Martinus Nijhoff;

Fant, G.C.M. (1967) The Nature of Distinctive Features, in: "To Honor Roman Jakobson", The Hague, Mouton & Co.;

Gimson, A.C. (1962) An Introduction to the Pronunciation of English, London, Matthew Arnold;

Jakobson, R. and Halle, M. (1956) Fundamentals of Language, The Hague, Mouton & Co.;

Kim, Chin-W. (1965) On the Autonomy of the Tensity Feature in Stop Classification (with special reference to Korean Stops), Word 21, 339/359;

Lisker, L. and Abramson, A.S. (1964) A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements, word 20, 384/422;

Moulton, W.G. (1962) The Vowels of Dutch: Phonetic and Distributional Classes, Lingua 11;

Slis, J.H. and Cohen, A. (1969) On the complex regulating the voiced-voiceless distinction, Lang.a.Sp. 12, 80/102 and 137/155;

Stetson, R.H. (1951) Motor Phonetics, Amsterdam, North-Holland Publishing Co.