(ARIPUC 15, 1981 p. 77-102)

ASPIRATED STOP CONSONANTS BEFORE LOW VOWELS, A PROBLEM OF DELIMITATION, - ITS CAUSES AND CONSEQUENCES

ELI FISCHER-JØRGENSEN
AND
BIRGIT HUTTERS

Attention is drawn to the fact that low vowels may show a gradual start after aspirated consonants, so that it is possible to consider the vowel to start at different points: (a) at voicing start, (b) at the start of Formant 1, (c) at the start of higher formants. The choice of delimitation is shown to have serious consequences for the statement of various temporal relations. It is argued that the choice of point (c) is the most rational choice, both from the point of view of production and perception. It also gives the most regular temporal relations.

I. INTRODUCTION

The study of temporal relations in speech has been given much attention lately, one reason being that these relations may give new insight into the way speech is organized from the point of view of motor commands, and perhaps also from the point of view of perception. But a statement of the temporal relations presupposes a relatively unambiguous delimitation of the units studied. This delimitation is often undertaken on the basis of acoustic curves, although in most cases the phonetician is more interested in an interpretation in terms of speech production and speech perception. Particularly in complicated consonant clusters the perceptual segmentation may be at variance with the articulatory boundaries. The delimitation of intervocalic aspirated consonants is relatively simple in this respect, since here articulatory and perceptual boundaries may be expected to coincide. But it raises a problem of a different kind which is caused by the gradual transition between aspiration and vowel, often found in the case of low vowels. The differences between a delimitation made on the basis of start of periodicity and a delimitation made on the basis of start of full formant structure of the vowel may be up to 60 ms in individual cases and more than 40 ms for averages, more than enough to have crucial consequences for the temporal relations set up on this basis. The problem cannot thus be brushed aside as negligible.

II. THE START OF THE CONSONANT

The determination of the start of a stop consonant does not give rise to serious problems. The only difficulty is that the implosion noise is generally too weak to show up on spectrograms except in the case of apicals. In this latter case it is often seen as a thin line; and as it appears immediately after the cessation of the formants of the preceding vowel phoneticians generally agree to define the start of the consonant as the end of the vowel formants in all cases. This means that the (very short) implosion is included in the measurement of the closure phase of the stop.

On oscillograms the corresponding point is the start of an abrupt decrease of the amplitude of the speech wave, coinciding with an abrupt fall of the intensity curve and a sharp drop of the fundamental frequency (Fo) curve. The oscillogram very often shows 10-30 ms of weak oscillations after this point.

Some Swedish phoneticians seem to have a slightly deviating tradition on this point. They often place the delimitation line between vowel and following voiceless stop at the point where voicing ends (e.g. Lindblom and Rapp 1972, Leanderson and Lindblom 1972, Löfqvist 1975a). The reason for this rule seems to be that in Swedish the abduction of the vocal folds often starts before the oral closure, so that there may be a short period of preaspiration, particularly after stressed vowels. In some cases the voicing does continue after the closure, but it is generally too weak to show up in oscillograms, cf. Fant (1973 (1969), p. 120), Karlsson and Nord (1970), Lindqvist (1972), with tracings of glottograms, and Löfqvist (1975b). It is not quite clear how Lindqvist determines the point of occlusion, but Karlsson and Nord have used electrodes on the lips, and they demonstrate that for four Swedish subjects the closure of the lips comes immediately after the cessation of voicing when the preceding vowel is unstressed, and often 10-40 ms later when the preceding vowel is stressed. -- However, in curves of Danish, German, and English stops, and also in Dutch and French, the normal case seems to be a continuation of voicing during a few centiseconds after the oral closure, and in Danish the glottal abduction seems to start immediately after the implosion. The appropriate point of delimitation must thus be the point where the intensity falls abruptly.

In velar consonants the decrease is often more gradual than in labials and dentals, but the area of uncertainty is rarely more than 10 ms.

III. THE ASPIRATION

A. THE PROBLEM

The release of a stop can be located safely on almost all types of curves.

On acoustic curves, particularly spectrograms, it is often possible to distinguish the transient noise of the release from a following fricative phase, which in its turn can be distinguished from the aspiration proper, characterized by a more *h*-like noise (see, e.g., Fant 1973 (1969), p. 111). On mingograms this three-way distinction cannot be made, and generally the three phases of transient noise, fricative noise, and aspiration are taken as one segment, which is sometimes called 'burst' (although this term is also used for transient plus fricative phase alone), or 'open interval', or simply 'aspiration' (thus used here in a wide sense). These finer segmentations will not be treated here. The problem we want to take up is the delimitation between aspiration and the following vowel.

In most cases the start of the vowel after a voiceless consonant does not raise any problems. Generally voicing starts simultaneously and abruptly in the whole spectrum. This is the case after fricative consonants, after unaspirated stops, and in most cases also after aspirated stops followed by high vowels. But in the case of low vowels there often seems to be a more gradual transition between aspiration and vowel, which may give serious problems of delimitation.

B. MATERIAL AND SUBJECTS

In order to throw some light on these problems we have examined a large number of curves of Danish words with stop consonants. In Danish *bdg* are voiceless and *ptk* strongly aspirated, for older speakers around 70 ms, for younger Copenhagen speakers around 90-100 ms (see Fischer-Jørgensen 1980). The investigation is based on recordings by 16 different Danish speakers who had acted as subjects in earlier investigations, some of them in more than one: A, the material used by Hutters (1979) (5 subjects: HU, MF, FJ, LG, and PA); B, part of the material used by Fischer-Jørgensen (1980), particularly lists M2 (6 subjects: HU, MF, LG, BM, JJ, and PM) and S5 (3 subjects: NK, BL, KS); C, the material used by Jeel (1975), which was re-measured for this purpose (6 subjects: LG, BM, VJ, EC, EH, and JJ); D, a material used by Reinholt Petersen for measurements of fundamental frequency and spoken by himself (NR); and finally E, a recording of *pale* and *pile* by PD.

Some subjects from material B were left out because of various problems, e.g. JR, who showed a variable vowel start, often with gradual appearance of F1, and BH, whose higher formants were so weak that their starting point could not be identified. PM has only been included for aspiration measurements, not for vowel durations, because the delimitation from the following consonant was uncertain.

The measurements of material A are based on a comparison between spectrograms and mingograms. As for material B, the measurements of list S5 are based on spectrograms, those of M2 on mingograms controlled by means of a restricted number of spectrograms. For material C, D, and E only mingograms were available. The mingograms contained a duplex oscillogram, a high fidelity intensity curve, an intensity curve highpass filtered at 500 Hz, and a fundamental frequency curve (material A also comprised photo-electrical glottograms). The mingograms did not contain any intensity curve highpass filtered at higher frequencies.

The slow start of the vowel is found more often in Danish *pa* and *ka* than after *t* which is affricated. *t* has therefore been left out, and only words with *pa* and *ka* have been examined and compared to words beginning with *pi*, *ki*, *ba*, *ga*, *bi*, *gi* and *fa*, *fi*. Danish *i* is very high, also when it is short; Danish *a* is a front vowel varying between cardinal vowel No. 4 and 3, for younger speakers closer to No. 3, thus [æ].

C. TYPES OF VOWEL START

An inspection of the curves showed a high degree of individual variation, but four main types could be distinguished: (1) some speakers have an abrupt start of all formants in both high and low vowels (although there is sometimes a strong fundamental and a certain weakening of F1 at the start). The abrupt start is mainly found in older subjects with relatively short aspirations (KS, NK, BL), but in one recording also in the case of JJ (list C), who has long aspirations. (In a different recording (list M2) he shows later start of higher formants.) (2) Some speakers start out with a few centiseconds of very low frequency vibrations (narrow band spectrograms show that the main intensity is in the fundamental), but there may also be some very weak striation in the region of the first formant; then, after 20-40 ms, all formants appear with full intensity (subjects LG and BM). (3) Some speakers start with a relatively strong first formant, but the higher formants do not appear until some centiseconds later (PA and JJ, list M2). (4) Finally, some speakers (in our material only female speakers: HU, MF, and FJ) often have a still more stepwise start: first low frequency (sometimes with very weak striations in the region of the first formant), then a strong F1, and finally the higher formants. In the following we will indicate the start of low frequency vibrations as point (a), the start of F1 as point (b), and the start of higher formants as point (c). There are, however, intermediate cases and it is sometimes difficult to decide when the first formant should be considered to start.

In oscillograms the low frequency start shows up as simple vibrations, often - but not always - of relatively low amplitude. The duplex oscillogram shows lack of higher components. The start of F1 and higher formants cannot be distinguished in the oscillogram. It is, however, possible to identify the start of the higher formants by means of an intensity curve, high-pass filtered at a higher frequency, e.g. 1500 Hz. Moreover, in Danish - at least in Copenhagen speech where stressed vowels generally have a falling fundamental frequency contour (which is particularly clear in low vowels) - there will be a turning point from rising to falling fundamental frequency at the point where the higher formants start. This has been controlled for several speakers and seems to be quite consistent. Some speakers do not always have the rising-falling fundamental frequency, but when they have it there is coincidence with the start of higher formants. Thus the turning point in the fundamental frequency curve could be used to identify point (c) in the cases where only mingograms were available.

Figure 1 shows some typical cases, and table 1 gives a survey of the magnitude of the differences in ms between points (a), (b), and (c). The difference a-c could be measured in all cases; (b) may coincide with (a) (PA, JJ) or with (c) (LG, BM), or it may not be measurable on the available curves (list C and HU, MF in list B). The average difference in ms between points (a) and (c) is 30.1 ms in *pa* and 24.5 ms in *ka*.

Apart from individual differences, other factors may influence the start of the vibrations. The force of the speaking voice plays a role. The higher the loudness level, the higher the chance of finding an early start of low frequency vibrations, and in this case they may even show up before high vowels (this has been shown by Preben Andersen (1981)). Some speakers often have a slightly later start of F2 and rising fundamental frequency at the start in high vowels. The preceding vowel may also be of some influence: we have, e.g., found a slightly earlier start of the vibrations in *api* than in *ipi*. Moreover, a quicker tempo favours voicing, and in quick connected speech an unstressed aspirated Danish consonant may be completely voiced. Finally, it cannot be excluded that the distance to the microphone may have a certain influence on the intensity of higher formants. It cannot, however, explain the different types of vowel start found in the present investigation. - The difference we have found between high and low vowels is thus fully valid only in the usual types of test sentences like "they said *pile*, they said *pale*", etc. - i.e., in distinct speech at a moderate tempo and loudness level.

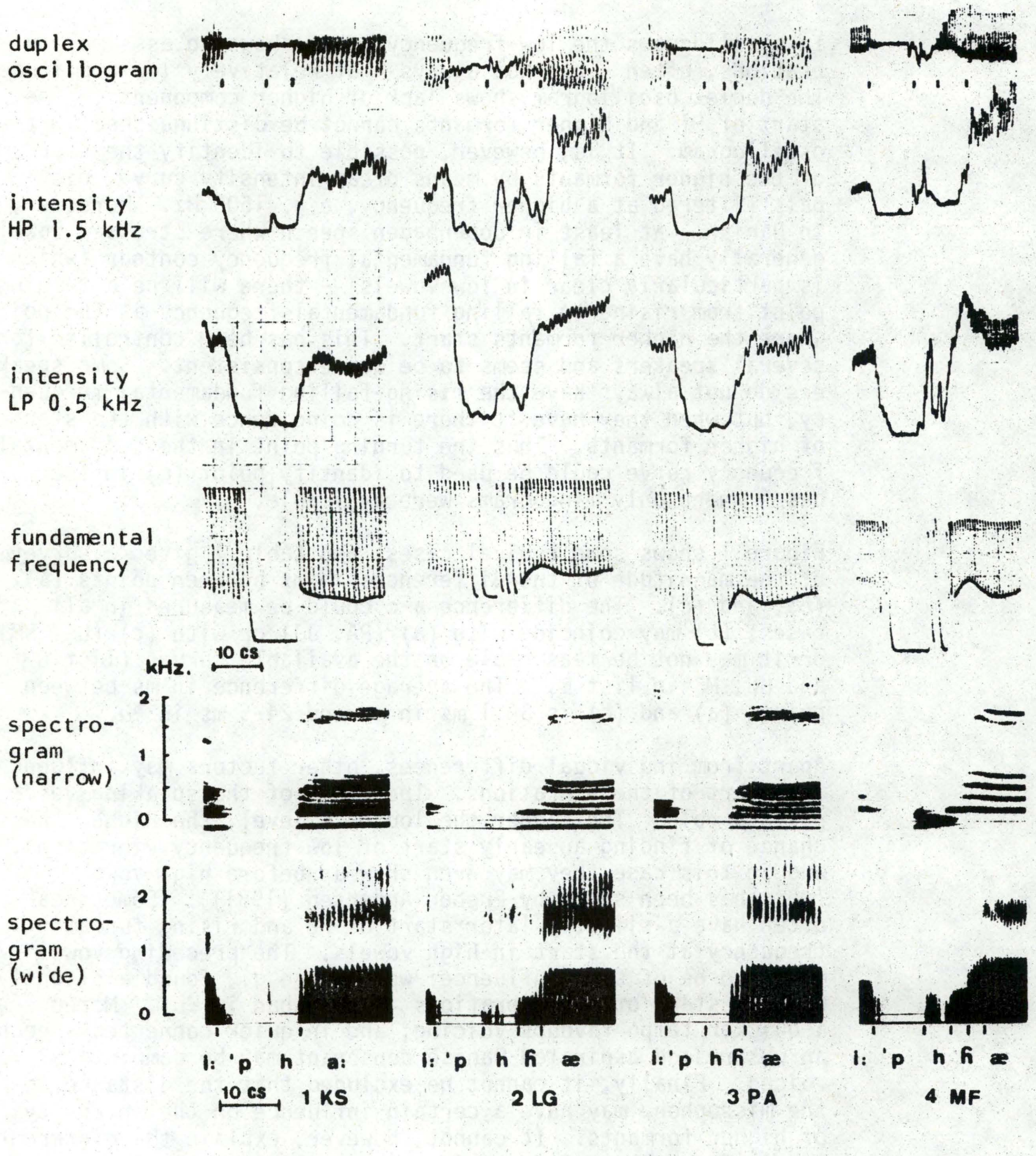


Figure 1

Four types of vowel start in Danish *pa*: 1 (KS) abrupt start; 2 (LG) start Fo-F1,F2 (a-bc); 3 (PA) start Fo,F1-F2 (ab-c); 4 (MF) start Fo-F1-F2 (a-b-c). (In MF's mingogram there is a click from a camera simultaneously with *p*'s open interval.)

Table 1

Differences between measurements in ms according to voicing start (point a), start of F1 (point b), and start of higher formants (point c) for speakers with gradual vowel start in *pa* and *ka*, separated according to the types mentioned in the text: a-b-c (type 4), a-bc (type 2), ab-c (type 3), a-c (cases in which point b could not be identified). (N = number of averages, each including 9 tokens, on the average. In the other tables, N = number of tokens.)

Type	Speaker	Material	N	*pa* Difference a-b	*pa* Difference b-c	*pa* Difference a-c	*ka* Difference a-b	*ka* Difference b-c	*ka* Difference a-c
a-b-c (4)	HU	A	1	17	29	46	7	28	35
	MF	A	1	17	16	33	19	10	29
	FJ	A	1	21	14	35	17	11	28
a-bc (2)	LG	B	3	34		34	36		36
	BM	B	2	32		32	24		24
ab-c (3)	PA	A	1		34	34		21	21
	JJ	B	1		21	21		30	30
a-c	HU	B	1			31			27
	MF	B	1			22			12
	EC	C	1			31			19
	EH	C	1			21			9
	VJ	C	1			12			16
	NR	D	1			34			32
	PD	E	1			36			
Grand means						30.1			24.5

D. CHOICE OF CRITERIA IN THE PHONETIC LITERATURE

The gradual start of the vowel may be particularly prominent in Danish with its long aspirations, but it is not restricted to Danish. We have found similar gradual transitions between aspiration and vowel in curves of Indian aspirated consonants (although not as a regular phenomenon) and in curves of German and English aspirated stops.

It is, therefore, not surprising that phoneticians who have measured vowel duration or duration of aspiration have used different criteria of delimitation. A few phoneticians consider the vowel to start at the release of the stop consonant (or at the end of the fricative phase). The main argument for this delimitation is that the supraglottal part of the vocal tract may be in the position for the vowel already at this time, and the formants - at least the higher formants - can often be traced in the aspiration noise; thus the vowel may be considered to have started, it is only voicing that begins later. This might be an acceptable point of view if the voiced part of the vowel were correspondingly shortened so that the sum of the aspiration (= voiceless vowel) and the voiced vowel would be of approximately the same duration as a vowel after unaspirated consonants. It is true that there is, e.g. in Danish, a certain shortening of the (voiced) vowel after aspirated consonants compared to unaspirated consonants, but this shortening is of only about 15-20 ms (sometimes in the case of speakers with very long aspirations 25 ms), and this modest shortening does not by any means compensate for aspirations of about 100 ms. If the aspiration were included as part of the vowel, short vowels after *ptk* might be acoustically more than twice as long as short vowels after *bdg*, and they would often be longer than long vowels after these latter consonants. This would give very irregular vowel durations. It can thus be concluded that considering the aspiration as part of the vowel in languages with long aspirations will lead to absurd temporal relations. (These arguments may not be valid for sonorant consonants following aspirated stops, but this problem will not be treated here.) Moreover, neither the speaker nor the hearer is aware of any difference in vowel duration after *ptk* and *bdg*. The vowels seem to be intended and perceived as being of the same duration. Obviously, the vowel, as the syllabic peak, has to be voiced in order to fill its function. - The remaining possibilities are thus points (a), (b), and (c).

Those who have measured oscillograms have generally considered the start of the voicing (point a) as the start of the vowel (e.g. Löfqvist (1975a), Kent and Moll (1969), Zue (1976)). This is the point which can be located with the highest degree of precision on oscillograms. The term "voice onset time", introduced by Lisker and Abramson, also at first glance seems to indicate a delimitation according to voicing, but it appears quite clearly from their discussion of segmentation problems (1964, p. 416-418, and 1967), that they do not include what they call weak "edge vibrations", appearing near the base line of the spectrogram, which they have found in a small number of utterances (their example is *k* + *a*). These edge vibrations are excluded because perceptual experiments have shown them to be inaudible. This means that in the cases of gradual transitions between aspiration and vowel they would not place the end of the voicing lag period at point (a) but rather at point (b) (start of the first formant) where voicing will probably be fully audible, but hardly at point (c). -- On the whole, those who base their measurements on spectrograms will hardly include these "edge vibrations" since they do not always show up very clearly in spectrograms (they may be partly covered by the base line and sometimes by the effect of puffs of air in the microphone after an aspirated *p*).

Lehiste and Peterson (1960) also indicate that they consider the vowel to start "at the moment in time at which periodic striations are stable in the first formant frequency" (thus point (b)). Still others choose point (c), thus Klatt (1975, p. 687) who states that "the VOT is indicated by the sudden onset of vertical striations in the second and higher formants" and Weismer (1979). Finally, it should be said expressly that placing the VOT limit e.g. at point (b) or (c) does not necessarily imply that one considers the vowel to start at this point (VOT means "(audible) voice onset", not "vowel onset"). Abramson (personal communication) is more inclined to consider the aspiration as a (voiceless) part of the vowel.

In earlier writings by one of the authors (FJ) there has been a certain vacillation between point (b) and point (c), whereas Hutters (1979) has chosen point (c).

E. DISCUSSION OF THE PRODUCTION AND PERCEPTION ASPECTS

As long as the only aim is acoustic measurements, the problem is not so important. One may simply measure the different acoustic segments. However, as mentioned above, one generally wants to make measurements which are meaningful from the point of view of production and/or perception. Moreover, when the purpose is to set up rules of temporal relations, one may prefer a delimitation of the units involved which permits a maximal generalization of rules, covering, e.g., both high and low vowels by the same rule. There is finally the practical problem of choosing a point which can be located with a sufficient degree of certainty on different types of curves.

Point (a) is the easiest one to locate on oscillograms. But this choice can be criticized both from the point of view of perception and production. From the point of view of production it is evident that vibrations starting in point (a) are produced before the glottis is closed for the vowel. That the vocal folds can vibrate in a relatively open position could be seen in the early high-speed films of glottal movements, e.g.

the Bell film and the film by Elizabeth Uldall, "Vocal Cord Action in Speech: A High Speed Study", 1957. That this is also the case in Danish *pa*- and *ka*-syllables has been demonstrated by means of a large number of simultaneous acoustic and glottographic recordings of Danish stop consonants (Hutters 1979). In *pi* and *ki* voicing normally starts very shortly before or at the end of the glottal gesture, whereas in *pa* and *ka* it starts at some point during the closing phase, sometimes even at the top of the glottographic curve. This difference can be explained by the fact that the open vocal tract in *pa* [phæ] and *ka* [khæ] causes a quick pressure drop above the glottis which permits voicing to start early, whereas the more constricted vocal tract in *pi* and *ki* will delay the intraoral pressure drop, and more so in *ipi* than in *api* because the tongue will be raised all through *ipi*. After Danish *t* the pressure drop is delayed by the affrication both before *i* and *a*. The earlier start of the voicing in very loud speech can be explained by the higher subglottal pressure.

When voicing starts before the glottis is closed the glottal spectrum will contain only relatively low components, and the formants will not appear until later, or only in the form of noise. As long as the vocal folds are separated the fundamental will also be relatively low, and it will be rising until the glottis is (almost) closed. This means that point (a) may occur at any moment of the glottal closing movement and before the tongue is in position for the vowel, depending on subglottal pressure, tempo, and other factors which do not have a direct relation to the segmentation of the speech chain in consonants and vowels but belong to the feature of voicing. Voicing is not a sufficient criterion for vowel start. Point (c), on the other hand, i.e. full formant structure and start of the Fo movement characteristic of the vowel, can be shown to coincide with the end of the glottal movement (see Hutters 1979). This will, moreover, give the same criterion for high and low vowels. It is thus a well defined point (see figure 2).

Point (a) can also be criticized from the perceptual point of view. As mentioned above, Lisker and Abramson found that these low frequency "edge vibrations" were not perceived as voicing, and in an experiment with tape cutting and splicing of Danish stop consonants (Fischer-Jørgensen 1972a) it was shown that the inclusion or removal of a segment of 10-15 ms containing such low frequency vibrations (and probably some weak high frequency noise which was not, however, visible on the spectrogram although the vibrations were) was decisive for the perception of the preceding stop as unaspirated or aspirated, thus this segment was heard as belonging to the consonant. Moreover, it can be stated more informally (we have not made any tests) that aspirations which, according to criterion (a) are of about 30-40 ms but according to criterion (c) about 80-90 ms, are heard as normal Danish aspirations.

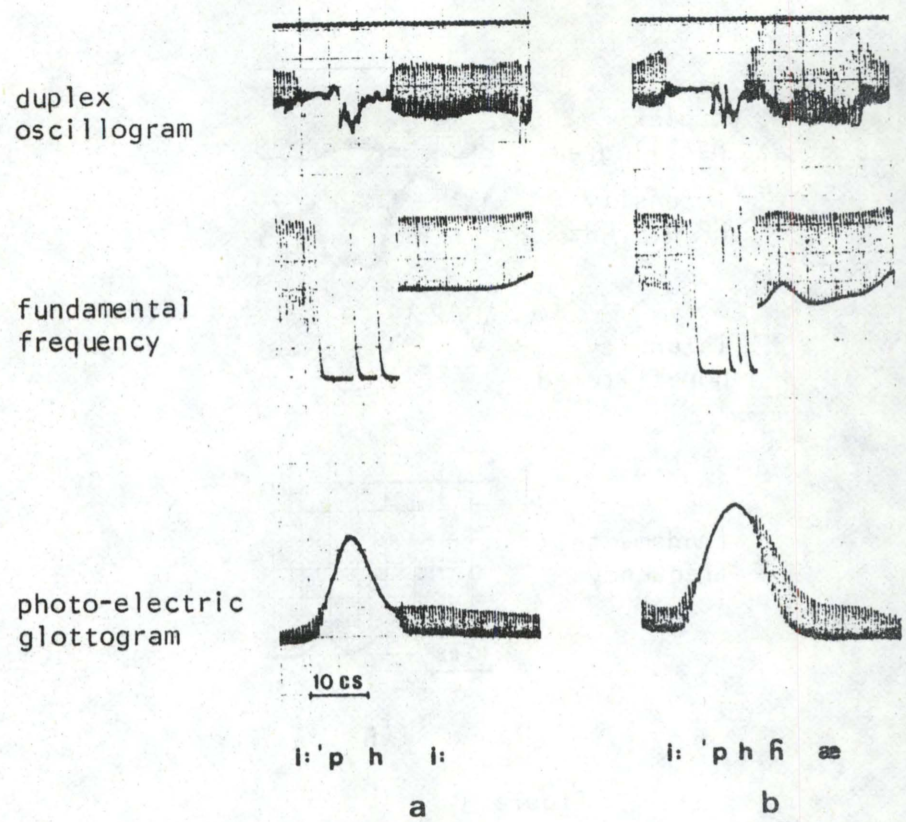


Figure 2

Danish *ipi* and *ipa*, subject MF

The rejection of point (a), i.e. start of periodicity, is corroborated by an investigation of Darwin and Pearson (1980). They have tested the *p-b* boundary on various continua of synthetic sounds and found that periodicity (in this case weak periodicity of formants) is of little importance for the perceptual boundary between *p* and *b*. Perceptually aspiration seems to end at a point where the formants of the vowel have reached a certain level of intensity. Probably this point coincides more or less with our point (c): start of higher formants. Further experiments are required to test whether overall intensity or intensity of particular formants is the determining factor.

Point (b) is less well defined than (c). From the point of view of production it is situated somewhere during the closing movement of the glottis, and acoustically the first formant has appeared as voiced, but the higher formants only as noise. Perceptual tests are needed on this point. Moreover, point (b) is rather difficult to locate on the oscillograms because there is often a gradual transition from low frequency vibrations to

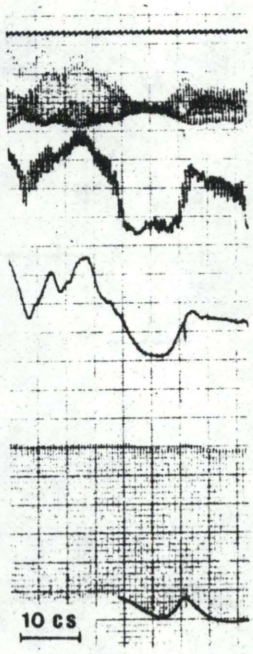


Figure 3

Danish voiced *h* in -*ha*-, subject EH

full first formant, and the intensity of the vibrations may be gradually increasing.

An argument supporting the choice of point (c) is that this point corresponds to the most appropriate point of delimitation between voiced *h* and a following vowel. Voiced *h* has a dip in intensity and fundamental frequency, and if the following vowel has a falling fundamental frequency, the turning point is very clear. Point (c) will give a common criterion for vowel start after aspirated voiceless stops and *h* (see figure 3).

F. CONSEQUENCES FOR THE DESCRIPTION OF TEMPORAL RELATIONS

Important arguments against point (a) and for point (c) can also be found by looking at the consequences of the choice for the regularity of rules concerning temporal relations.

The relations examined are: (1) differences of vowel duration after *ptk* compared to *bdg* and *fs*; (2) differences of duration between the vowels *a* and *i*; (3) differences of duration of aspiration before *a* and *i*.

1. DIFFERENCES OF VOWEL DURATION AFTER *ptk* COMPARED TO *bdg*

As mentioned above, Danish vowels are generally slightly longer after *bdg* than after *ptk*. This can be shown for high vowels where no problem of delimitation arises. Since the materials were not originally recorded for this purpose, there are not always examples of all consonants, but the number is sufficient for a thorough examination.

The difference between the durations of *i* after *ptk* and *bdg* could be measured for 15 different speakers, some of them taking part in various recordings. There were 54 averages, each based on about 9 tokens, and they all showed the vowel to be longer after *bdg* than after *ptk*, on the average 16.4 ms. This is thus a very well established relation. Finally, at least for some subjects, the choice of point (c) gives a smaller range of variation.

The relations for the vowel *i* can now serve as a standard of comparison for measurements of the vowel *a* after *ptk* and *bdg* according to different criteria of delimitation. In table 2a and figure 4a the averages of *i* and *a* are compared for four speakers who have an abrupt start of both vowels. The grand mean of the difference between the durations of the vowels after *b* and *p* is 22.0 ms for *i* and 21.6 ms for *a*. For *g*-*k* it is 16.2 ms for *a* (there are no comparable cases with *i*). Table 2b and 2c and figure 4b-c show the differences for speakers with gradual vowel start. In table 2b and figure 4b the difference between the durations of *a* after *pk* and *bg*, delimitated after two different criteria, (a) and (c), is compared to the durations of *i* after *pk* and *bg* for the same speakers. The grand mean for *i* is 17.4 ms; for *a* measured from the start of the vibrations (criterion (a)) it is -0.4 ms, whereas measured from the start of the higher formants (corresponding to the turning point of the pitch curve) it is 24.1 ms. Table 2c and figure 4c show the difference between the duration of *a* after *pk* and *bg* according to criteria (a) and (c) for those speakers who did not speak words with *i*. The grand mean for point (a) is -6.8 ms, for point (c) 18.3 ms.

These comparisons show quite clearly that there is full agreement between the measures for *a* and *i* if the vowel *a* is defined as starting at point (c), but absolute disagreement if it is defined as starting at the beginning of voicing (point (a)). In this latter case the vowel *a* will be of approximately the same duration after *pk* and *bg* or, for some speakers, clearly shorter after *bg*, in contradistinction to the consistently longer duration after *bg* for the vowel *i*, which was also found for *a* in the case of the speakers with abrupt vowel start. The relations are also valid for the individual speakers except that in table 2b criterion (a) is preferable to (c) for JJ (*b*-*p*), and (a) and (c) are approximately equivalent for BM (*b*-*p*). For LG (a) is preferable both in the case of *b*-*p* and *g*-*k* when compared to the values for *i* of the same speaker but not when compared to the differences for other speakers, and in table 2c criterion (c) is obviously preferable for LG.

Table 2

Differences of vowel duration after *pk* and *bg* (in ms)

a. Speakers with abrupt start of all vowels

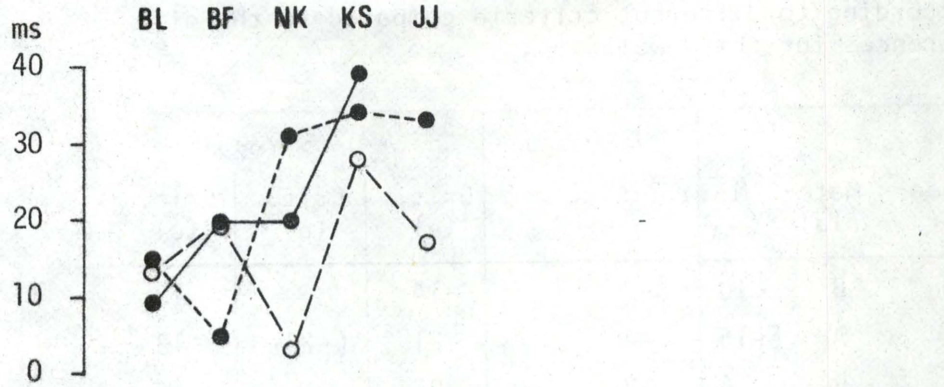
Spea-ker	Mate-rial	N	*(b)i>* *(p)i*	N	*(b)a>* *(p)a*	N	*(g)a>* *(k)a*
BL	B	2	9	6	15	6	13
BF	B	2	20	2	5	2	20
NK	B	2	20	6	21	6	3
KS	B	2	39	6	34	6	28
JJ	C			20	33	20	17
Grand means			22.0		21.6	18.9	16.2

b. Speakers with gradual vowel start in *pk* + *a*

					Crit. (a) (PA and JJ: ab)		(Crit. (c)	
Spea-ker	Mate-rial	N	*(b)i>* *(p)i>*	*(g)i>* *(k)i>*	*(b)a>* *(p)a*	*(g)a>* *(k)a*	*(b)a>* *(p)a*	*(g)a>* *(k)a*
LG	B	10	10	7	4	6	26	28
BM	B	10	9	17	-3	1	24	27
HU	B	6	8	24	-20	-7	11	20
PA	A	5-15	27	15	-19	-12	15	9
MF	B	10	26	23	12	0	33	12
JJ	B	10	11	31	24	9	45	39
Grand means			15.2	19.5	-0.3	-0.5	25.7	22.5
			17.4		-0.4		24.1	

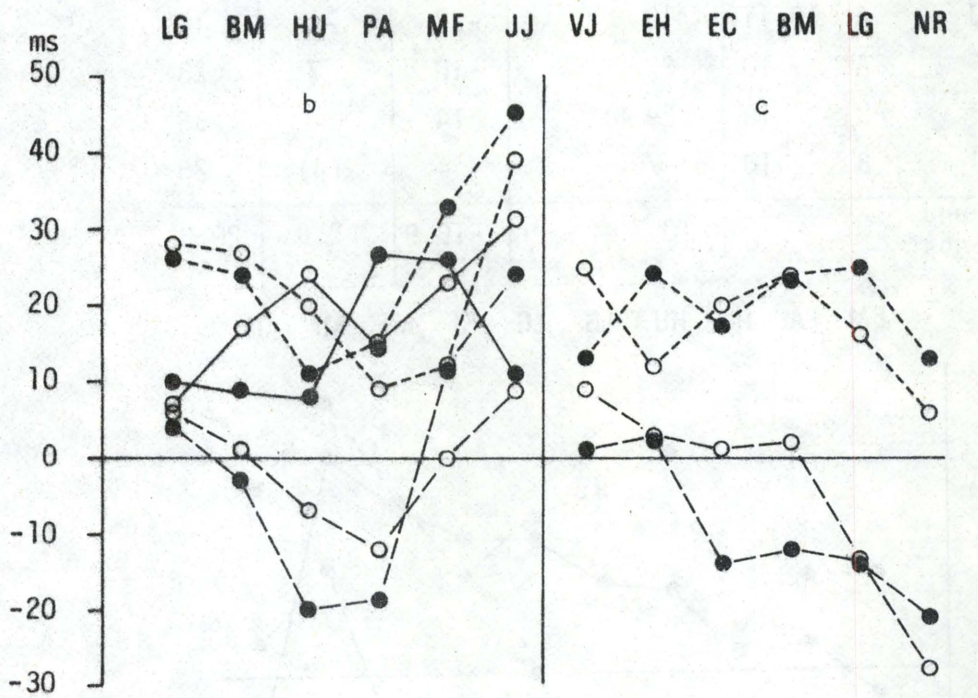
c. Speakers with gradual vowel start in *pk* + *a* (no examples with *i*)

		N		Crit. (a)		Crit. (c)	
Spea-ker	Mate-rial	*b-p*	*g-k*	*(b)a>* *(p)a*	*(g)a>* *(k)a*	*(b)a>* *(p)a*	*(g)a>* *(k)a*
VJ	C	10	5	1	9	13	25
EH	C	12	10	3	3	24	12
EC	C	12	8	-14	1	17	20
BM	C	11	10	-12	2	24	24
LG	C	8	9	-13	-13	25	16
NR	D	9	9	-21	-28	13	6
Grand means				-9.3	-4.3	19.3	17.3
				-6.8		18.3	



a. Speakers with abrupt vowel start in *a*.

●——● *(b)i* > *(p)i* ●---● *(b)a* > *(p)a* ○——○ *(g)a* > *(k)a*



b - c. Speakers with gradual vowel start in *a*. (LG and BM have spoken two different materials.)

●——● *(b)i* > *(p)i* ○——○ *(g)i* > *(k)i*

●—–● *(b)a* > *(p)a* (crit.a) ●----● *(b)a* > *(p)a* (crit.c)

○—–○ *(g)a* > *(k)a* (crit.a) ○----○ *(g)a* > *(k)a* (crit.c)

Figure 4

Differences of vowel duration after *pk* and *bg*.

Table 3

Differences of vowel duration (in ms) after *p* and *f*. Speakers with gradual vowel start in *pa*, delimitation according to different criteria compared to the differences for the vowel *i*

Speaker	Material	N	*(f)i > (p)i*	*(f)a > (p)a* Crit. (a)	Crit. (b)	Crit. (c)
BM	B	10	-3	-14		13
PA	A	5-15	4	-21	(-21)	13
HU	B	6	3	-25		6
HU	A	10-11	11	-17	0	29
LG	A	10-13	12	-18		24
LG	B	10	18	0		21
FJ	A	10-11	18	-22	-1	13
MF	A	10	22	-10	7	23
MF	B	10	29	14		36
JJ	B	10	-7	4	(4)	25
Grand means			10.7	-10.9	2.0	20.3

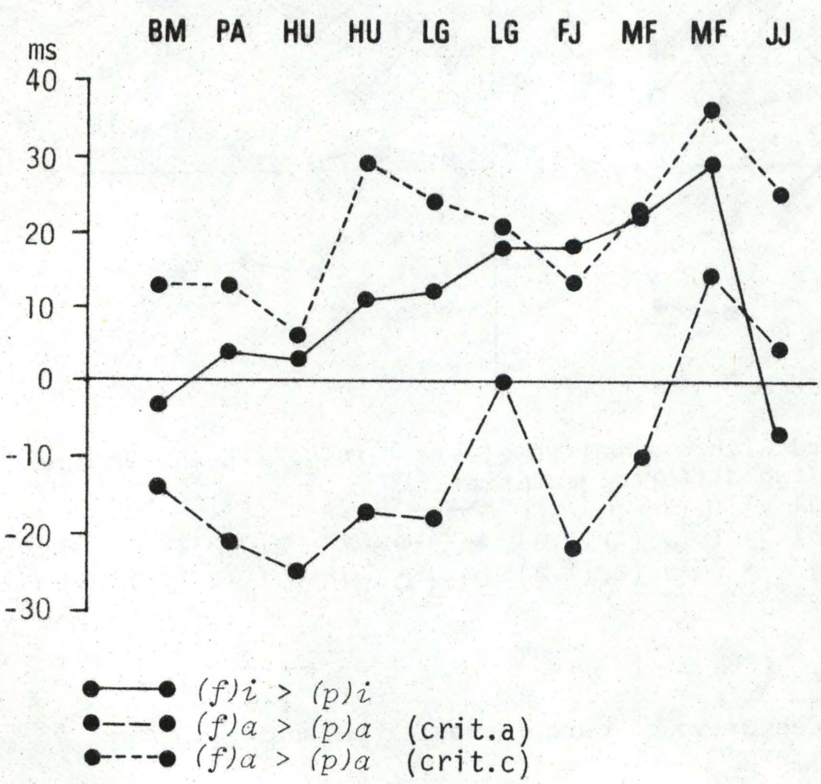


Figure 5

Differences of vowel duration after *f* and *p*. (LG, MF, and HU have spoken two different materials.)

Since the measurements of material B, C and D were made on the basis of mingograms (with a few spectrograms for control), it was not possible to distinguish the start of F1. However, in the case of PA and JJ, spectrograms were taken of all examples. They showed, as mentioned above, that for these two subjects F1 starts at once, i.e. point (a) and (b) coincide. For PA it is evident that (c) is preferable to (b) as defining the start of the vowel, whereas for JJ this is only obvious for *g-k*. Subject BH, who was left out because her higher formants were so weak that point (c) could not be identified with certainty, had a large difference between (a) and (b), and the differences according to (b) were in agreement with her differences for the vowel *i*.

The results concerning the relations between vowels after *ptk* and *bdg* can be supported by a comparison of vowel durations after *p* and *f*. High vowels are longer after *f* than after *p*, although the differences are not quite as regular as in the case of vowels after *ptk* and *bdg*. There are no examples for the speakers with abrupt vowel start. Table 3 and figure 5 show that for the speakers with gradual vowel start the differences between the durations of *a* after *f* and *p* is in better agreement with the differences for *i* when the measurements are made according to criterion (c) than when they are made according to criterion (a), the grand mean for *i* being 10.7 ms and for *a* 20.3 ms according to criterion (c) but -10.9 ms according to criterion (a) (thus giving a shorter vowel after *f* than after *p*). This is also valid for all individual speakers except BM and JJ. It has been possible to distinguish points (a), (b), and (c) for subjects HU, MF and FJ. Point (c) is obviously preferable when the values are compared to those for the vowel *i*. For PA (a) and (b) coincide, and the measurements have been listed both under criterion (a) and (b). In his case (c) is also in better agreement with the measurements for *i*.

2. DIFFERENCES OF DURATIONS BETWEEN THE VOWELS *a* AND *i*

It is a universal finding that the vowel *a* is longer than *i*. This is also true of the present material, and the difference is so large that it exists irrespective of the criteria of delimitation chosen. But it is possible to compare the magnitude of the difference according to different criteria with the magnitude of the difference in the positions where the delimitation does not give any problems. Table 4 shows that the difference between *a* and *i* is around 50 ms in these latter cases, i.e. for speakers who have abrupt vowel start in *a* (table 4a, grand mean 50.7 ms), similarly in vowels after *g* and *s* for all speakers (table 4b, grand mean 59.2 and 51.4 ms after *g* and *s*, respectively), and finally after *b d g f s* for PA (table 4b, grand mean 48.0 ms). After the consonant *k* criterion (c) gives approximately the same result (55.8 ms, for PA 56.0 ms), but with a relatively large range of variation, whereas criteria (a) and (b) give considerably higher values (86.7 and 78.7 ms, respectively, and criteria (ab) for PA gives 74 ms).

Table 4

Differences between the duration of the vowels *i* and *a*

a. Speakers with abrupt vowel start in both vowels

Speaker	Material	N	*a* > *i* after *p*	*a* > *i* after *t*	*a* > *i* after *k*
BF	B	2			45
NK	B	12	46	48	40
BL	B	12	55	54	47
KS	B	12	67	59	46
Grand mean				50.7	

b. Speakers with gradual vowel start in *ka*

Speaker	Material	N	*a* > *i* after *g*	after *s*	after *k* Crit. (a)	after *k* Crit. (b)	after *k* Crit. (c)
HU	A	10-11		47	96	89	61
MF	A	10-12		52	77	58	48
FJ	A	10-11		57	106	89	78
LG	A	10		65	112		63
HU	B	6	82	39	113		66
MF	B	10	62	34	85		73
LG	B	10	62	50	63		41
BM	B	10	43	52	59		33
JJ	B	10	47	67	69	(69)	39
Grand means			59.2	51.4	86.7	78.7	55.8

				a > *i* Crit. (ab)	*a* > *i* Crit. (c)
PA	A	5-15	after *b:* 43 *d:* 46 *g:* 32 *f:* 65 *s:* 54	after *p:* 89 *t:* 73 *k:* 60	after *p:* 59 *t:* 70 *k:* 39
Grand means			48.0	74.0	56.0

This is also valid for the individual speakers, except for BM and LG in material B (but in material A, LG shows the same relations as the other subjects). The results for JJ are ambiguous. On the whole, this relation thus supports the choice of criterion (c).

3. DIFFERENCES IN THE DURATION OF THE OPEN INTERVAL OF THE STOPS BEFORE *i* AND *a*

There is a universal tendency for the open interval after stop consonants to be longer before high vowels than before low vowels. This may be explained by the slower escape of air after the release because of the narrower constriction (but in a low back [ɑ] the constriction in the pharynx may be expected to have a similar effect). The consonant *t* (which is not included here) always has a longer open interval before *i* than before *a* [æ], whereas for *p* and *k* the open interval is only consistently longer before *i* if the start of voicing (point (a)) is taken as the start of the vowel. However, as shown in table 5 and figure 6, the differences obtained by choosing this point (30.4 ms for *p* and 30.5 ms for *k*) are, on the average, considerably higher than the difference found for those speakers who have abrupt vowel start (9.0 ms for *p* and 12.2 ms for *k*). Point (b), too, gives higher values (25.3 and 22.0 ms). On the other hand, point (c) results in a deviation in the opposite direction so that there is, on the average, practically no difference before *i* and *a*. As shown clearly in figure 6, point (c) gives negative values (i.e., the aspiration is shorter before *i* than before *a*) for half the subjects (here point (a) is better) and positive values for the other half (here point (a) gives too large differences compared to the subjects with abrupt vowel start).

The question is now how consistent one should expect these relations to be in the case of aspirated *p* and *k* before *i* and *a*. If the tendency is due to the slower escape of air after the release because of a more constricted vocal tract, one should expect the differences to be particularly clear in cases of short open intervals, e.g. for Danish voiceless *bdg* and French unaspirated *ptk*, and this is indeed the case (see, e.g., Fischer-Jørgensen 1972b and 1980). In Danish *bdg* and, according to various investigations, e.g. Kagaya and Hirose (1975), in Indian unaspirated *ptk*, the glottis is almost closed at the release and ready for voicing, and the duration of the open interval is thus exclusively dependent on the time it takes for the pressure to decrease in the oral cavity. Vowel start is equal to voicing start. But in aspirated stops the constriction evidently influences the start of voicing (cf. the large differences between the open intervals before *a* and *i* when the delimitation is made at voicing start (point (a)). But the intraoral pressure may have decreased to zero when the adduction movement of the vocal cords is finished and all formants appear, so that one should not necessarily expect a difference at point (c). In Fischer-Jørgensen 1980 a consistent difference was found for *ki-ka*, but in some cases the delimitation

Table 5

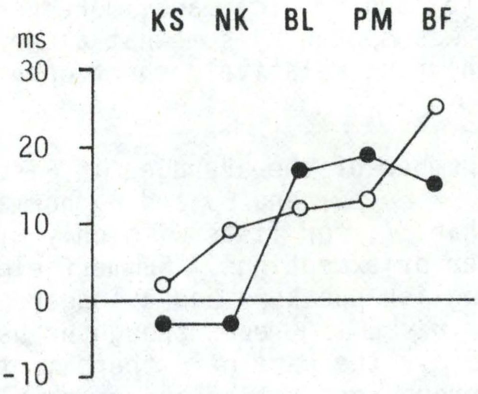
Differences between the open intervals of *p* and *k* before *a* and *i* (in ms)

a. Speakers with abrupt vowel start

Spea-ker	Mate-rial	N	*(p)h(i)>* *(p)h(a)*	*(k)h(i)>* *(k)h(a)*
KS	B	12	-3	2
NK	B	12	-3	9
BL	B	12	17	12
PM	B	10	19	13
BF	B	2	15	25
Grand means			9.0	12.2

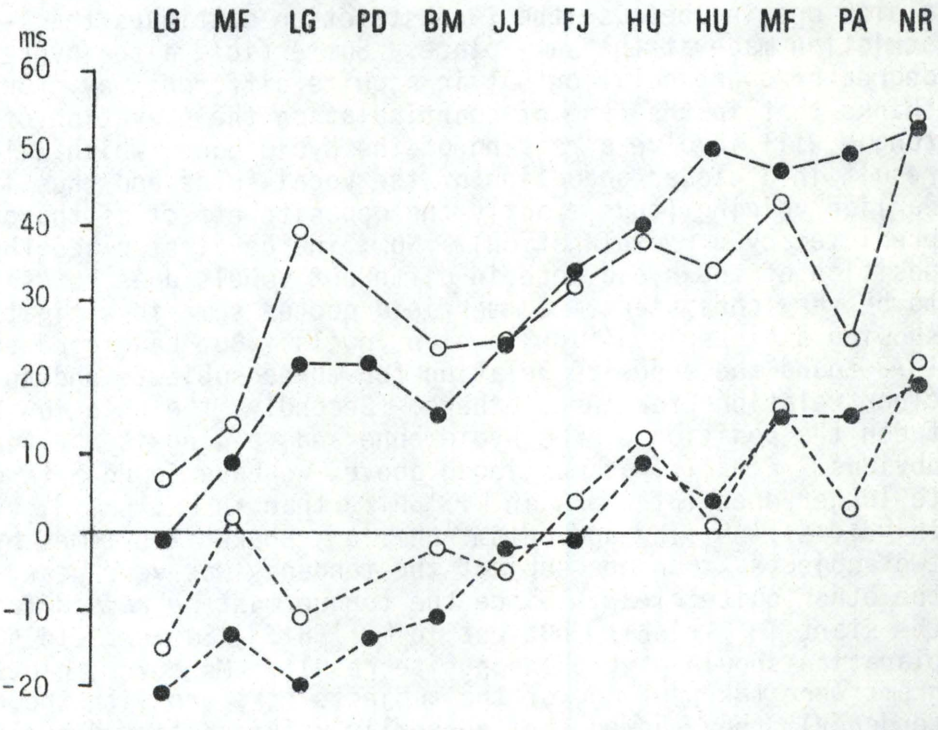
b. Speakers with gradual vowel start

			(p)h(i)> *(p)h(a)*			*(k)h(i)>* *(k)h(a)*		
Spea-ker	Mate-rial	N	Crit. (a)	Crit. (b)	Crit. (c)	Crit. (a)	Crit. (b)	Crit. (c)
LG	B	10	-1		-21	7		-15
MF	B	10	9		-13	14		2
LG	A	10-13	22		-20	39		-11
PD	E	10	22		-14			
BM	B	10	15		-11	24		-2
JJ	B	10	25	(25)	-2	25	(25)	-5
FJ	A	10	34	13	-1	32	15	4
HU	B	6	40		9	38		12
HU	A	10-11	50	33	4	34	27	1
MF	A	10	47	30	15	43	24	16
PA	A	5-15	49	(49)	15	25	(25)	3
NR	D	9	53		19	54		22
Grand means			30.4	25.3	-1.6	30.5	22.0	2.5



a. Speakers with abrupt vowel start.

●——● *(p)h(i)* > *(p)h(a)*   ○——○ *(k)h(i)* > *(k)h(a)*



b. Speakers with gradual vowel start - according to criteria (a) and (c). (LG, MF, and HU have spoken two different materials.)

●— — ● *(p)h(i)* > *(p)h(a)* (crit.a)
○— — ○ *(k)h(i)* > *(k)h(a)* (crit.a)
●- - - -● *(p)h(i)* > *(p)h(a)* (crit.c)
○- - - -○ *(k)h(i)* > *(k)h(a)* (crit.c)

Figure 6

Differences between the open intervals of *p* and *k* before *i* and *a*.

may have been made at point (b) and not (c), and, moreover, a large part of the material was spoken by somewhat older subjects or non-Copenhageners who have relatively short open intervals in aspirated stops.

As far as *p* is concerned, a number of the averages in Fischer-Jørgensen 1980 were based on *pi py pu*, and *py* and *pu* normally have longer open intervals than *pi*; for lists with only *pi* (e.g. M2), there were a number of exceptions. Summerfield 1975 found that 5 out of 6 English speakers had a longer open interval in *pa* [pʰɑː] than in *pi* [pʰiː], even though he used voicing start as a criterion. In the case of *p* there is the special condition that the tongue need not be raised completely at the release (particularly not in the case of a diphthongized British [ɩi] and, on the other hand, a British [ɑː] may have a very narrow constriction in the pharynx). The open interval in *pi* (and also in [pʰɑː] and [kʰɑː]) may thus depend partly on the degree of coarticulation, whereas *ki* always has a slow opening because the *i*-constriction continues the *k*-constriction made at the same place. Summerfield also invokes the degree of coarticulation but in a quite different way. He thinks that in the case of coarticulation the elevation of the tongue will involve a raising of the hyoid bone, which will result in a closer adduction of the vocal folds and thus in earlier voicing (thus exactly the opposite effect of the one predicted by our explanation). Now, in the first place the position of the hyoid bone in different vowels does not seem to be very consistent. Summerfield quotes some investigations showing a higher position in high vowels. But Ladefoged et al. 1972 found the opposite relation for three subjects and no clear relations for three others. Secondly, the relation between the position of the hyoid bone and voicing is not very obvious. Finally, as mentioned above, we have found a tendency to longer open intervals in [iːˈphiː] than in [æːˈphiː], viz. in [divisiːˈphiːlə] and [disæːˈphiːlə], spoken ten times by two subjects. For one subject the tendency was very weak, for the other quite clear. Since the tongue must be raised from the start in [iːˈphiː] but not in [æːˈphiː], Summerfield's explanation should give the opposite result. Moreover, glottograms were taken of one of the subjects (the one with the weak tendency); they showed that generally voicing started earlier during the adduction of the vocal folds after *a* than after *i*. This has strengthened our belief that the main factor in the delay of voicing is the constriction in the vocal tract. This is also supported by the fact that Danish *pu* and *py*, in which the lip opening remains small in the transition from *p* to the vowel, have a longer open interval than *pi*. And it is further corroborated by inspection of intraoral pressure curves. Generally, voicing does not start until the pressure curve has reached zero level. In *pa* this may happen almost instantaneously, in *pi* the decay is slower, but variable, and in *pu* it often takes up to 70-80 ms, which means that the slow decay may really delay voicing in aspirated stops in some surroundings, and even sometimes vowel start according to criterion (c), which presupposes both start of voicing and start of higher formants. But as it need not do so in *pi*, nor, in the case of long aspirations in *ki*, we do not think that the lack of difference between the open intervals in *pi-pa* and *ki-ka* can be used as an argument against the choice of criterion (c).

4. DIFFERENCES IN INTRINSIC Fo

It should finally be mentioned briefly that the choice of delimitation criterion may also influence the relations of intrinsic Fo. Generally, the Fo curve starts at a higher frequency value after Danish *ptk* than after *bdg*, but since the curve rises from (a) to (c) in *pa-*, *ta-*, *ka-*syllables, the difference will be larger in point (c) than in point (a). Jeel 1975 used point (c), and for one of her subjects the relations got reversed if point (a) was chosen.

IV. CONCLUSION

The purpose of the present paper was in the first place to draw attention to a problem of delimitation which is not of a negligible magnitude (up to 45 ms for averages, and up to 60 ms in individual cases). All three criteria ((a) start of periodicity, (b) start of F1, and (c) start of higher formants) have been used by different phoneticians, and we should like to emphasize that the so-called "weak edge vibrations" may show up with quite considerable amplitude in oscillograms so that point (a) may seem to be a natural point of delimitation both for vowel start and (still more obviously) for VOT, when oscillograms are used as a basis. We have further shown that different points of delimitation result in rather different temporal relations, and we have argued for using point (c). It seems quite evident to us that point (c) is preferable to point (a), (1) because it is physiologically better determined as coinciding approximately with the closure of the glottis, (2) because the low frequency vibrations starting in point (a) are perceptually weak and are not heard as part of the vowel, and (3) because point (c) gives much more regular temporal relations. The evidence for preferring point (c) (start of higher formants) to point (b) (start of F1) does not carry the same weight because it is based on a much more restricted material, but in almost all cases point (b) gives less regular temporal relations than point (c) and, moreover, it is rather difficult to locate on oscillograms.

REFERENCES

Andersen, P. 1981: "The effect of increased speaking rate and (intended) loudness on the glottal behaviour in stop consonant production as exemplified by Danish *p*", *Ann. Rep. Inst. Phon. Univ. Cph.*, this issue

Darwin, C.J. and Pearson, M. 1980: *What tells us that voicing has onset?*, unpublished paper given to the Institute of Acoustics Speech Group, York, December 1980

Fant, G. 1969: "Stops in CV-syllables", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn., Stockholm 4/1969*, and *Speech Sounds and Features, 1973*, p. 110-139

Fischer-Jørgensen, E. 1972a: "Perceptual studies of Danish stop consonants. Tape cutting experiments with Danish stop consonants in initial position", *Ann. Rep. Inst. Phon. Univ. Cph. 6*, p. 104-168

Fischer-Jørgensen, E. 1972b: "PTK et BDG français en position intervocalique accentuée", *Papers in Linguistics and Phonetics for the Memory of Pierre Delattre*, (The Hague), p. 143-200

Fischer-Jørgensen, E. 1979: "Temporal relations in consonant-vowel syllables with stop consonants based on Danish material", in *Frontiers of Speech Communication Research* (Lindblom, B. and Öhman, S. eds.), (London)

Fischer-Jørgensen, E. 1980: "Temporal relations in Danish tautosyllabic CV sequences with stop consonants", *Ann. Rep. Inst. Phon. Univ. Cph. 14*, p. 207-262 (revised and enlarged version of 1979)

Hutters, B. 1979: *Glottisfunktionen ved produktion af ustemte obstruenter i dansk, en glottografisk og fiberoptisk undersøgelse*, unpublished report (Institute of Phonetics, University of Copenhagen)

Jeel, V. 1975: "An investigation of the fundamental frequency of vowels after various consonants, in particular stop consonants", *Ann. Rep. Inst. Phon. Univ. Cph. 9*, p. 191-212

Kagaya, R.A. and Hirose, H. 1975: "Fiberoptic, electromyographic, and acoustic analyses of Hindi stop consonants", *Annual Bulletin, University of Tokyo, Logopedics and Phoniatrics 9*, p. 27-46

Karlsson, I. and Nord, L. 1970: "A new method of recording occlusion applied to the study of Swedish stops", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn., Stockholm 2-3/1970*, p. 8-18

Kent, R.D. and Moll, K.L. 1969: "Vocal tract characteristics of stop cognates", *J. Acoust. Soc. Am. 46*, p. 1549-1555

Klatt, D.H. 1975: "Voice onset time, frication and aspiration in word initial consonant clusters", *J. Speech and Hearing Research 18*, p. 686-706

Ladefoged, P., DeClerk, J., Lindau, M., and Papçun, G. 1972: "An auditory-motor theory of speech production", *Working Papers in Phonetics, UCLA 22*, p. 48-75

Leanderson, R. and Lindblom, B.E.F. 1972: "Muscle activation for labial speech gestures", *Acta Otolaryngologica 73*, p. 362-373

Lehiste, I. and Peterson, G. 1960: "Studies of syllable nuclei", *Speech Research Laboratory Report 4*, (Ann Arbor)

Lindblom, B. and Rapp, K. 1972: "Reexamining the compensatory adjustment of vowel duration in Swedish words", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn., Stockholm 4/1971*, p. 19-25

Lindqvist, J. 1972: "Laryngeal articulations studied on Swedish subjects", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn., Stockholm 2-3/1972*, p. 10-27

Lisker, L. and Abramson, A.S. 1964: "A cross-language study of voicing in initial stops: acoustical measurements", *Word 20*, p. 384-422

Lisker, L. and Abramson, A.S. 1967: "Some effects of context on voice onset time in English stops", *Language and Speech 10*, p. 1-29

Löfqvist, A. 1975a: "On intrinsic and extrinsic Fo variations in Swedish tonal accents", *Phonetica 31*, p. 228-247

Löfqvist, A. 1975b: "A study of subglottal pressure during the production of Swedish stops", *J. Phonetics 3*, p. 175-189

Summerfield, Q. 1975: "Aerodynamics versus mechanics in the control of voicing onset in consonant-vowel syllables", *Speech Perception, Series 2, Number 4*, (Belfast), p. 61-72

Uldall, E. 1957: *Vocal cord action in speech: a high-speed study*, (film), University of Edinburgh and Swiss Federal Institute of Technology

Weismer, G. 1979: "Sensitivity of voice-onset-time (VOT) measures to certain segmental features in speech production", *J. Phonetics 7*, p. 197-204

Zue, V.W. 1976: "Acoustic characteristics of stop consonants: a controlled study", *Technical Report 523*, Lincoln Laboratory, Massachusetts Institute of Technology