ON BRITISH ENGLISH p-, b-, sp- and -s + b-

Henry Petersen

1. Introduction

1.1. The problem

The problems connected with the clusters consisting of s + stops in English must be intriguing indeed, since so many linguists (and educationalists, for that matter) have done an almost incredible amount of work to solve them. The attention of the present writer was also engaged, and the result is the article below, which, it is hoped, will throw some light on the articulatory conditions of p-, b-, sp- and -s + b-.

One of the pedagogical questions facing Danish teachers and learners of English is whether our pronunciation of written sp- [sb̥] can be transferred to English without modification. Without unduly anticipating the conclusion of this article it may be said that the answer must be in the negative.

Otto Jespersen (1958, p. 51) states "that the English sounds [p], [t], [k] remain practically unaltered when preceded by [s]." His view is supported by Daniel Jones (1962, p. 139), who says that "Scandinavians also have a tendency to replace p by b̥ when it occurs ... after s as in spend ...".

A.C. Gimson (1964, p. 146) finds that "when /s/ precedes /p, t, k/ initially in a syllable, there is practically no aspiration, even when the syllable carries a strong accent, cf. pin [pʰIn] and spin [spIn] ..." and (not quite consistently, perhaps, p. 48) that "when /p, t, k/ follow an initial /s/, ... they are realized with no aspiration even when stressed". These phonetic (auditory?) facts are interpreted to the effect that, phonemically, the words spin, steam, scum may be trancribed either as /spIn, sti:m, skʌm/ or as /sbIn, sdi:m, sgʌm/ without ambiguity. Providing that Gimson's phonetic basis is sufficiently broad, the latter transcription is, phonologically, as sound as the former, of course.

The first incentive to inquire into these things came 5 or 6 years ago from a discussion with Professor Poul Steller, of Danmarks Lærerhøjskole. He maintained the views of Jespersen and Jones, perhaps with some modification. I agreed.

The second impetus was given by N. Davidsen-Nielsen (1969a 1969b), whose experimental procedures and the phonological results derived from the experiments appeared to me to be methodically unsound, involving, it would seem, some confusion of the phonetic and phonemic levels.

The achievement of the object outlined above rendered necessary not only an experimental investigation of the articulation of p-, b-, sp- and -s + b-, but also a statistical analysis of the varying lengths of these sounds.

1.2. Linguistic material

The linguistic material consisted of two groups of sentences called Ia, Ib, Ic, Id (40 sentences) and II, la; II, lb; II, lc; II, ld; II, 2a, etc. (440 sentences). The number 480 was not attained, however, partly owing to miscounts, partly on account of inferior quality of a few of the recordings.

Group I comprised four sentences with the intonation indicated, each spoken ten times in succession:

Ia I ˈswallow a ˈbitter ˎpill today
Ib I ˌswallow a ˋbitter ˌpill today
lc I ˌswallow a ˌbitter ˋpill today
ld I ˌswallow a ˌbitter ˏpill?

The object of the analysis of group I was to find out whether the realizations of p and b, respectively, in the same phonetic surroundings, but in different tonetic groups differ in respect of measurable characteristics to such a degree that they must be considered as belonging to the same/different statistical population(s). Furthermore - as in Frøkjær-Jensen et al. (1971) - there is an indication of the position in time of the plosion relatively to the glottal gesture, if any. In this way it was hoped that an exact assessment of the occurrence of aspiration, if any, especially in the sp- cluster, would be rendered possible.

Group II comprised 11 sub-groups, each of 4 sentences. The sounds under investigation were said within the frame of a carrier sentence, word-initially, and followed by the eleven "pure" English vowels, in the following order:

II, 1a	I say ˎpeak today	II, 1b	I say ˎspeak today
II, 1c	I say ˎbeak today	II, 1d	'Jack's 'beak is ˎred

The other words were: pill, spill, bill, Jack's bill (sub-group 2), pell, etc. (sub-group 3), pan (sub-group 4), par (sub-group 5), pore (sub-group 6) pot (sub-group 7), pun (sub-group 8), purr (sub-group 9), pull (sub-group 10); in this group was included the only nonsense word used: spull [spʊɫ]), poon (sub-group 11). Each sentence was spoken 10 times in succession in the order given above.

From Group II were selected 3 sub-groups, 1, 2 and 9, with the object of elucidating the articulation of p-, sp-, b- and s- + b- (114 sentences). Furthermore, sentences from sub-group 5 were selected and compared with sentences from sub-group 1 with a view to finding out whether in comparable phonetic and tonetic surroundings the mean length of consonants preceding vowels with maximum difference in the degree of opening is influenced significantly by this difference.

1.3. The subject

The subject was an Englishman, who gave the following account of his accent and educational background: 26 years of age, B. A. English from Cambridge. M. A. Applied Linguistics - Essex University. At present completing Ph. D. in linguistics.

Accent: Near R. P. Southern English imposed upon South Wales English. Very little remaining of previous dialect, except in the case of, e. g., orthographic flew, blew [flIu, blIu].

According to this information the subject - hereafter referred to as P. M. - is assumed to be representative of his linguistic community. Obviously, the results of investigations of such limited scope as the present experiment are in principle of highly limited validity because the material offers no possibility of estimating several important factors. It is hoped, nevertheless, that inferences drawn from the sentences recorded may have a certain general application as regards the (social) group of speakers of R. P.

1.4. Instrumentation and comments on the set-up

The output of the Fabre Glottograph was tapped with a time constant of 4 seconds. The input to the transpitchmeter (settings: LP filter 120 Hz, HP filter 70 Hz) was taken from the Fabre Glottograph, consequently the output curves were physiological. The transducer of the Photoglottograph was inserted through the subject's nose and fixed in the oesophagus. An aerometer was placed before the mouth, and a Revox Recorder (tape speed 7.5"/sec.) was connected to the microphone while Group II was being recorded.

The speed of the recording paper was in all cases preset at 10 cm/sec. As, however, the speed of the mingograph was not exactly 10 cm/sec., but 10,25 cm/sec., all figures are divided by 1.025, the results thus obtained being very near approximations to the true time.

No doubt one would be to some extent justified in reviving the objections of the classical phoneticians against experimental phonetics (Eli Fischer-Jørgensen (1962), p. 115 and (1957), p. 435). Though P. M. as a subject was all that could be desired, unaffected psychologically, calm and patient, even his speech was influenced by the apparatus to such an extent that it cannot be said to be natural. This is particularly the case with the nasals, which sound as if at the time of recording the subject had a cold in his head.

1.5. Introductory remarks on the results from group I

p and b were chosen as objects of investigation for the same reasons as adduced in the above-mentioned article by Frøkjær-Jensen et al.: "The study is limited to labials, partly because this simplifies the problems since there is no affrication present, and partly because the choice of labials minimizes the risk of artifacts in the glottogram stemming from tongue movements" (p. 123).

Artifacts occur, nevertheless, on account of the movements of the back of the tongue, most perceptibly in open back vowels. Here the articulatory movements may cut off the light from the Photoglottograph to such an extent that glottal gestures leave no traces in the mingogram. Besides, there is a very marked artifact in most of the n´s.

Both p and b are word-initial and intervocalic. Their surroundings are chosen so as to be very nearly identical, viz. allophones of [ə] and [I], whereas the intonation of the sentences varies. In these circumstances, such significant differences in length as might be established between the means of the four groups of b´s and the four groups of p´s respectively must be ascribable to varying stress and intonation.

This assumption is the basic idea of this article.

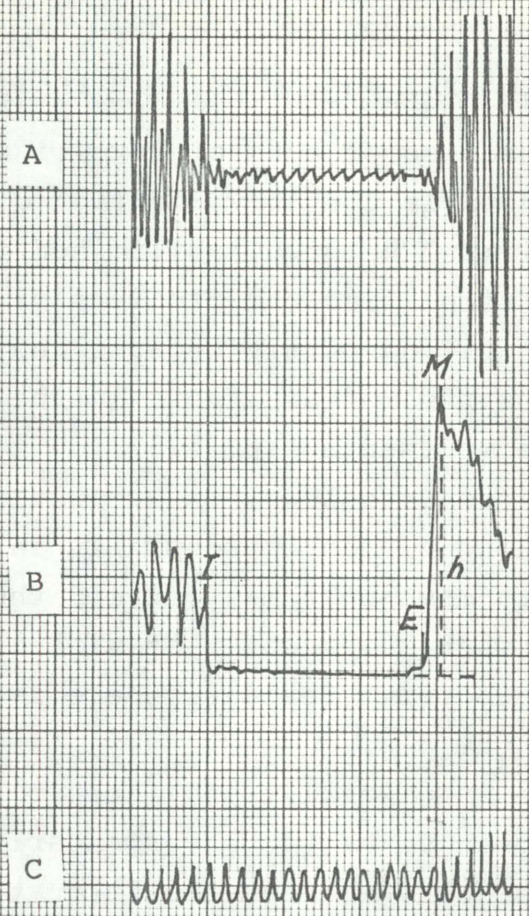
2. Measurements of b

2.1. Definition of the quantities measured in the sound b

Figure 1 illustrates the Photoglottograph and aerometer curves and the microphone oscillogram.

The "dip" of the curve between the points I and M is hereafter called the trough (also in p).

Fig. 1: parameters in b



A: Microphone oscillogram

B: Aerometer curve

C: Photoglottogram

It will be seen from Fig. 1 that the aerometer curve contains several salient points of interest for the measurement of the total duration of b. Point I is chosen at the vertex of the last well-defined vibration of the vocal chords in [ə]. Point M indicates the maximum deflection of the aerometer curve after the plosion. On close inspection of the material of group I it was found that in all partly devoiced specimens of b M coincided very neatly with the onset of vowellike vibrations in [I], whereas deviations were sometimes noted in fully voiced b's.

In the bulk of these cases there was perfect coincidence; but in some the onset of the vowel occurred from 1.5 cs to less than 0.5 cs before M.

Consequently, the total duration of b is measured along the time axis from I to M, and, when necessary, this measurement is adjusted in accordance with the distance between the vowel segments in the microphone oscillogram in order to obtain the best possible accuracy.

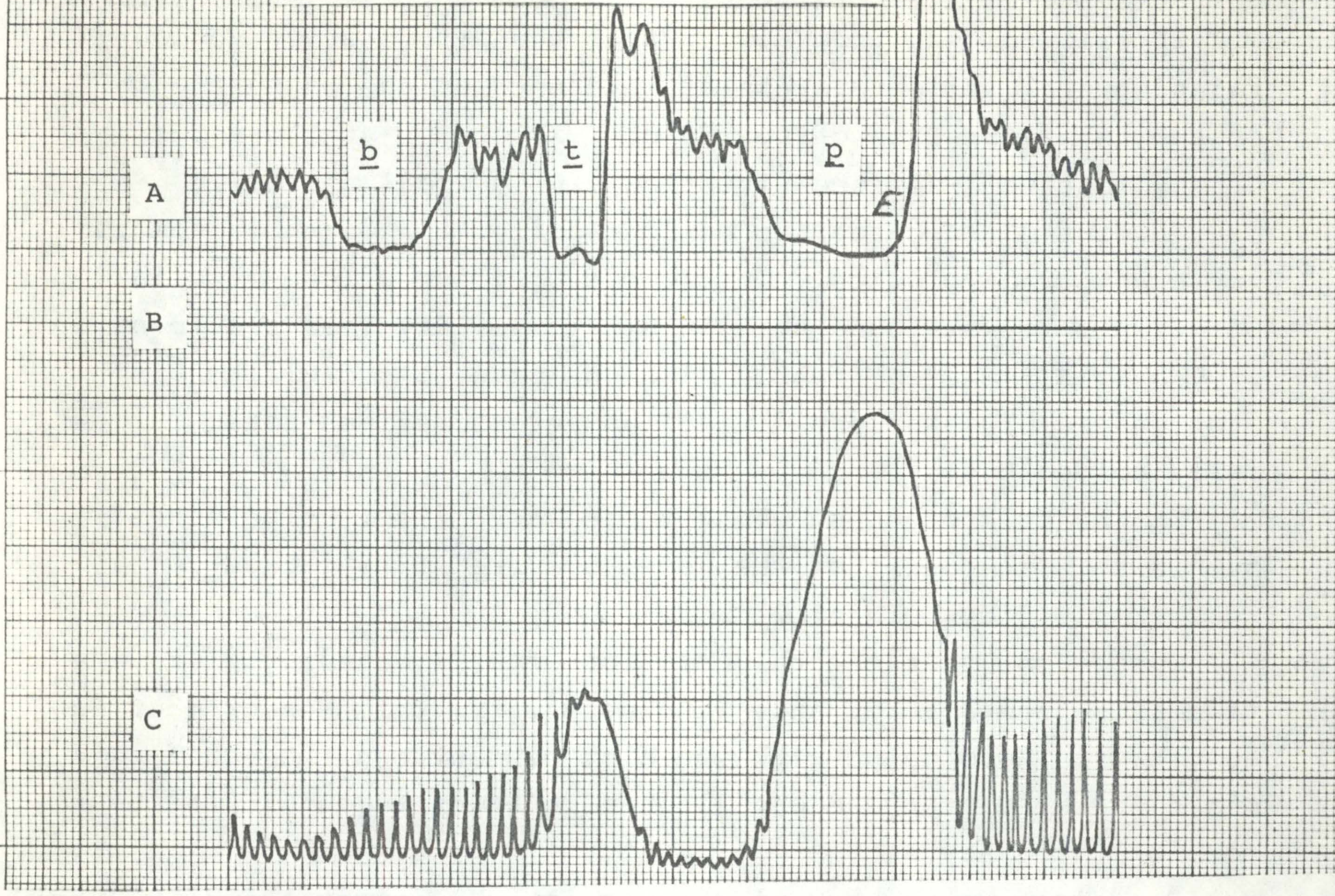
E indicates the time of the plosion, EM (along the time axis) the duration of the open interval, while IM - EM is an expression of the closure stage in a wide sense: transition + hold stage.

On account of the "floating" zero-line (see Figures 2a and 2b) the exact duration of EM cannot be calculated in most of the fully voiced b's, many of which seem to be fricative or "flapped" varieties of the sound (Gimson 1964, p. 154). It naturally follows that the mean values of EM given below are not representative.

In the partly devoiced b's and in some of the fully voiced ones the duration of EM was calculated as shown in Fig. 2b.

The ordinate h is a relative measure of the maximum airflow in M, measured in a cubic unit, e.g. cm^3/sec. Any ordinate between E and M will be a measure of the relative airflow at the given moment, whereas the slope of the curve indicates the increase in airflow at that moment. The slope of the curve is

Fig. 2a, showing a "flapped" variety of b with unmeasured open interval



A: Aerometer curve

B: Zero line in pauses between sentences

C: Photoglottogram

Fig. 2b, showing a partly devoiced b with measurable open interval

understood to be tan. of the angle between the abscissa and the tangent touching the curve at the point investigated. When, as was the case in the present experiment, the apparatus is not calibrated, the slope indicates the relative increase in airflow per unit of time.

As the curve between E and M may be taken to be roughly linear, it follows that its slope is indicative of the relative increase in airflow per second at any time of the interval EM.

This assumption underlies the following measurements of relative increase in airflow.

2.2. Variations in the duration of word initial b

As far as the duration of word-initial b is concerned, it is important to decide whether samples of b's drawn at random and in random numbers from uniform phonetic, but tonetically varying groups - like those established above - constitute a homogeneous material from which the average total duration of b could be calculated with confidence. This would not seem to be a likely proposition (Eli Fischer-Jørgensen 1970, p. 37 ff), and the figures below also suggest that the answer must be in the negative.

The average durations of b in cs in Ia through Id are:[1]

Ia	ˈbitter	9.4	*** (Ia–Ib)	* (Ia–Id)
Ib	ˋbitter	11.2		
Ic	ˌbitter	7.6	*? (Ic–Id)	
Id	ˌbitter	8.6		

The difference between the means of b in a stressed syllable with (relatively) high and static tone (Ia) and b in nuclear syllable with High Fall (Ib) is highly significant (exceeding the

1) One asterisk indicates significance at the 95% level (or better), two and three asterisks at the 99% and 99.5% levels respectively. o = not significant.

99.9% level), which must very nearly rule out the possibility that the difference is accidental. As, however, the two sounds occur in the same phonetic environment (with the possibly best approximation), it would seem safe to conclude that the difference in total length must be ascribed to membership of different tonetic groups (this does not answer the question whether b's in the same tonetic, but in different phonetic surroundings constitute distinct statistical populations).

The other two analyses (a b in a "normal" nuclear syllable does not, unfortunately, occur in the material) result in a similar, but perhaps a little more non-committal no to the question posed, the significance level in the case of Ia and Id exceeding the 95%, but not the 97.5% level, and in that of Ic and Id the t value being not quite, but very nearly equal to the t value of the 95 % level (this is suggested by the ? after the asterisk).

The conclusion of the above considerations must be that the duration of word-initial b in intervocalic surroundings cannot be calculated without due allowance being made for rhythmic and tonetic groupings.

2.3. Degree of voicing in b

There are rather large discrepancies between the amount of voicing shown by the Photoglottograph and the amount shown by the physiological duplex oscillogram (tapped from the Fabre Glottograph) - the latter showing longer intervals of voicelessness than the former.

It was therefore decided to measure the amount of devoicing from the duration of the glottal gesture in the Photoglottograph curve.

The figures are shown in table I below.

The average duration of all partly devoiced b's and the ratio of devoicing to total duration are given in parentheses, the average duration of all fully voiced items in square brackets.

As the numbers in question are too small, no conclusions can be drawn, but the figures may give interesting hints.

TABLE I

	fully voiced items	partly devoiced items	average duration of b in cs	average devoicing in cs	devoicing in %
Ia ˈbitter	3	6	9.4 (10.0) [8.4]	1.4	14.9 (20.8)
Ib ˋbitter	2	8	11.2 (11.3)[10.9]	2.4	21.4 (26.4)
Ic ˌbitter	7	2	7.6 (9.9) [7.0]	0.4	5.3 (19.3)
Id ˌbitter	5	5	8.6 (9.1) [8.0]	1.1	12.8 (24.1)

The above ratios suggest that the voicing of b is to some extent dependent on the total duration of the sound: the shorter it is, the smaller the amount of devoicing in it, and the greater the probability of its being fully voiced.

In group I all b's with a total duration of 8.4 cs or less were fully voiced.

Allowance being made for the small populations of the present material, the general conclusion would seem to be warranted that the duration of b is dependent on the rhythmic/ tonetic conditions of the words in which the sound occurs and also (as will appear later) on the length of the sentence.

Consequently, the maximum probability of intervocalic b being fully voiced will occur in weakly stressed words in (relatively) long sentences.

2.4. Survey of b in group I

TABLE II

		N	s	M in cs	open inter-val in cs	clo-sure in cs	h	V (slope)	tan. V
Ia	'bitter	9	1.1	9.4	2.2	7.7	12.5	68^{o}-84^{o}	5.9994
Ib	ˋbitter	10	1.2	11.2	2.4	9.0	14.1	71^{o}-84^{o}	6.5341
Ic	ˌbitter	9	1.7	7.6	2.3	7.3	10.2	69^{o}-83^{o}	4.3227
Id	ˌbitter	10	0.8	8.6	2.0	6.8	10.7	68^{o}-83^{o}	6.1893

(open interval and closure:) these figures are not representative (see 2.1.)

3. Measurements in p

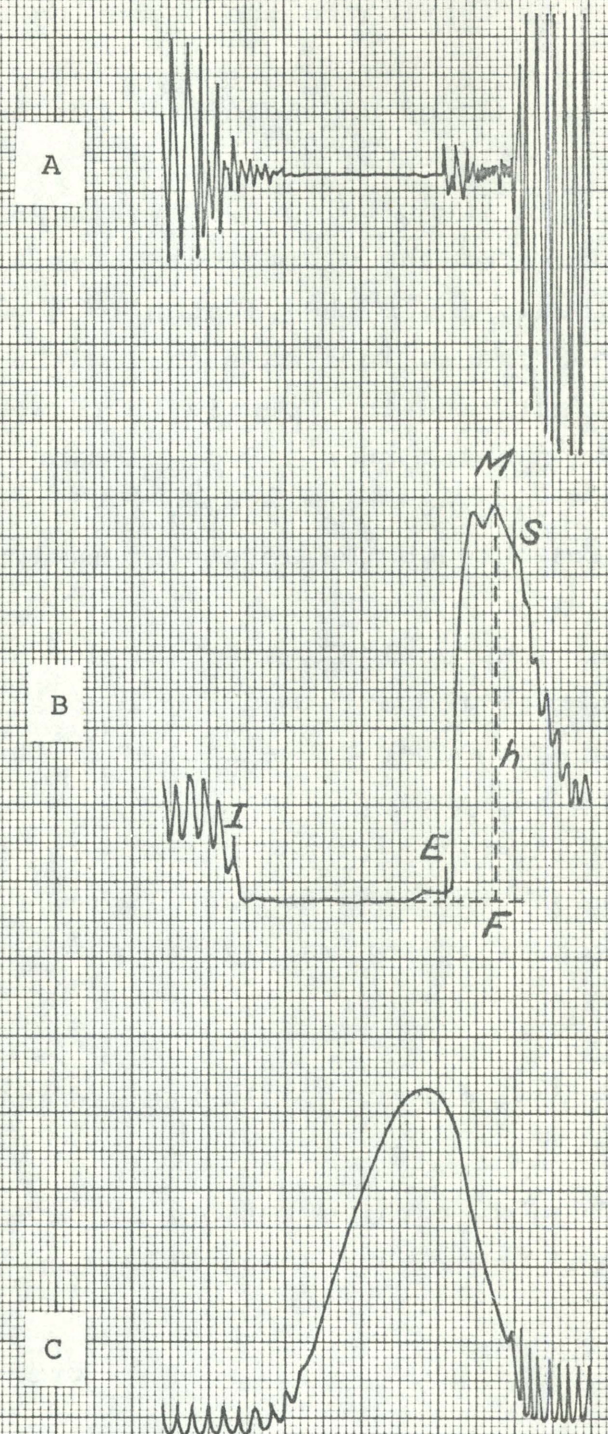
3.1. Definition of the quantities measured in the sound p

The parameters of the sound p are shown in Fig. 3.

The total duration of the sound p is measured along the time axis from I to S, and the measurement is compared with the corresponding distance between the vowel segments in the microphone oscillogram in order to obtain the best possible accuracy. Point I is placed at the vertex of the last well-defined vibration of the vocal chords in [ə] and S at the starting point of the first regular vibration of [I] after the aspiration phase (in the microphone oscillogram).

Very frequently point E - the lowest point in the aerometer curve - did not coincide with the plosion, so that in many cases the airflow through the instrument had started before the plosion. The cause of this is probably movements of the lips preceding the plosion proper.

Fig. 3: parameters in p



A: Microphone oscillogram

B: Aerometer curve

C: Photoglottogram

The open interval is measured in the microphone oscillogram from the onset of the aspiration to the beginning of vowellike vibrations.

The closure is calculated as the difference between total duration and open interval.

In groups Ic and Id the plosion nearly always took place during the opening phase of the glottal gesture (2 instances in Ic could not be decided), in Ib the plosion took place during the closing phase in 6 cases, in 1 case during the opening phase (3 could not be decided), and in group Ia there were 2 instances of plosion during the opening phase, 2 of plosion during the closing phase, and 3 could not be decided.

h is the ordinate through the aerometer maximum normal to the axis in F.

The angle V is an expression (generally somewhat "idealized") of the angle between the time axis and the longest nearly rectilineal segment of the aerometer curve from E to the point of maximum deflection of the curve.

As will be seen later, in (s)p-, where point E indicates the time of the plosion, the line segment IS (along the time axis) is in most cases of approximately the same length as IF, so that ES≈EF.

3.2. Variations in total duration, closure and open interval in word initial p

The average values of total duration, of closure and open interval (all measured in cs) are shown below.

```
               total                                open
               duration          closure            interval
-------------------------------------------------------------------------
Ia ˎpill      14.5\_______         10.4\_______       4.2\_________
                  >***    |            >O      |          >***  |    |
Ib ˈpill      13.2/    ***|  |     10.0/     **|  |   3.1/------O-- |
                  >***    |  O         >***    |  O       >O    |   O
Ic ˋpill      15.6/_______|  |     12.0/_______|  |   3.6/______|  *
                  >*         |         >*         |       >O        |
Id ˏpill      14.8/__________|     10.9/__________|   3.9/__________|
```

It will be seen that the total duration of intervocalic word-initial p is highly sensitive to shift of stress and tone. Also the closure of p is affected, but not to the same degree, perhaps.

The open interval seems to be the stablest component of the sound: only the short open interval of Ib (unstressed, low and level tone) differs significantly from Ia and Id. In all other cases there is no significant difference.

The figures seem to indicate that the open interval of p in nuclei with Low Fall and Low Rise is comparatively long. It may be a little surprising that the duration of the open interval in the High Fall group (Ic) does not differ significantly from that of the other groups.

The difference between M in Ia and Id is not significant, which means that p's in Low Fall and Low Rise groups constitute one statistical population. The difference in average duration may be supposed to be accidental and not attributable to membership of different tonetic groups. - N.B.! The length of the sentences is not the same, so that the validity of the conclusion may be contestable.

In all other cases the difference between the means is significant or highly significant, for which reason the groups must be considered as belonging to different statistical populations.

The conclusion of the above is that in uniform phonetic surroundings the total duration of p varies significantly with membership of different tonetic groups (with the said exception). The stronger the stress and the longer the tonetic glide, the greater the total duration.

On the whole, the standard deviations for the four groups of p's are smaller than those for the four groups of b's. Statistically, the duration of p must therefore be said to be a relatively more stable quantity than that of b. The reason is not far to seek. Phonetically, the b's are two different sounds:

a partly devoiced and statistically rather stable sound and a fully voiced and comparatively unstable sound, which in some cases, it would seem, may be "flapped" or even fricative. As an example of the unstableness of b it may be mentioned that the standard deviation for b in group Ic (ˌbitter) is 1.7 for the whole population; the range is 5.4 cs. The range of the fully voiced b's is 4.2 cs, of the partly devoiced items it is 0.9 cs (there are only 2 such sounds in the population). In p in Ib (ˌpill) the standard deviation is 0.7, and the range is 2.4 cs.

3.3. IS and IF values

It is of importance to remember that in initial p total duration (IS) is always distinct from the distance in time between the implosion and the moment of maximum airflow, so that IS > IF, whereas in p preceded by s the difference between the two values tends towards zero.

3.4. Increase in airflow on release of p (tan. V values)

It will be seen that the tan. V measurements show a surprising amount of uniformity, the four groups only differing a few degrees from one another. As some uncertainty is inherent in the method of measurement, a certain caution should be exercised in the interpretation of the differences between the tan. V values. No doubt, however, it would be safe to suppose that of the four groups Ic has the relatively greatest increase in airflow per unit of time, and that this quantity is dependent on the extent of the tonetic glide (and the stress?).

3.5. Comparisons between the lengths in groups I and II

Some statistical calculations were made in order to compare lengths of certain segments selected from groups I and II.

a. Ic total duration of p in ˋpill and II, 2a total duration of p in ˋpill

The difference between the two means is significant (exceeding the 99.5 % level). Both occur in intervocalic positions in High Fall groups, though not between the same vowels.

Alternative interpretations present themselves:

(i) the difference between the means is either ascribable to the fact that the phonetic surroundings of the two p's are not quite the same, [ə-p-I] in Ic and [ei-p-I] in II, 2 a, or to

(ii) the rhythmic-tonetic differences between the sentences.

Ic has 9 syllables, of which - besides the nucleus - at any rate two (Nos. 2 and 5) carry a certain amount of stress, whereas II, 2a has only one stress, viz. on the nuclear syllable.

It is hardly possible from the material to make out a strong case for the rejection of (i). However, two statistical calculations speak in favour of this.

Firstly, the difference between the total duration of p in II, 2a and in II, 3a is not significant (below the 60 % level): they belong to the same statistical population. These two sentences are rhythmically and tonetically identical, whereas phonetically they are "unilaterally" identical. Thus the difference in degree of opening/place of articulation for [I] and [e] does not bring about a significant difference between the p's in these phonetically different surroundings.

Secondly, the word pill in the sentences Ic and II, 2a was segmented into [p] and [Iɫ] on the aerometer curve (according to the usual criteria).

The mean of the total length of the latter segment was in Ic 19.9 cs, s = 2.1. The corresponding figures were for II, 2a 26.1 cs and 1.2. The difference between these means is highly significant (exceeding the 99.95 % level). Thus the two segments belong to separate statistical populations. As the phonetic surroundings of [Iɫ] in the two words are identical, there is every probability that the difference arises from tonetic conditions.

Moreover, the difference between the mean of the word pill (total length) in Ic (M = 35.5 cs, s = 2.0) and the mean of the word pill in II, 2a (M = 43.9 cs, s = 1.4) is also highly significant (exceeding the 99.95 % level).

On the strength of this evidence the first sentence of the alternative (i) is rejected, and the difference between the means is supposed to arise from rhythmic-tonetic "causes".

b. Possible influence of following vowel

Though the difference between the means of p in II, 2a and II, 3a is not significant, the possibility remains that significant differences between the means of the two sounds might be concomitant to greater differences in degree of opening/place of articulation of the vowels following p/b.

Comparisons were only made between total duration of p in II, la (peak) and of p in II, 5a (par) and between total duration of b of the same sub-groups (beak, bar). For the p's the difference between the means was highly significant (exceeding the 99.95 % level), and for the b's it was very probably significant (exceeding the 97.5, but not the 99 % level). As the rhythmic-tonetic conditions of these sentences are identical in pairs, it is hereafter taken for granted that at any rate extreme differences in degree of opening/place of articulation may influence the total length of p/b significantly. (N.B. perhaps the difference between open/closed syllable should not be disregarded, but further experiments are required to verify/falsify this proposition).

In consequence of the facts established above it was considered safer to make comparisons between items from within the same group and not between items from different groups.

Though the influence of the degree of opening/place of articulation has not been sufficiently clarified, there is, on the whole, no doubt that if reliable statistics of quantity in English are to be established, due allowance must be made not only for this problem, but also for rhythmic-tonetic phenomena.

Failing this, one runs the risk of uncritically confusing quantities from distinct statistical populations.

3.6 Survey of p in group I

TABLE III

	N	s	total duration in cs	open interval in cs	closure in cs	IF in cs	h	V (slope)	tan. v
Ia ˎpill	7	0.7	14.5	4.2	10.4	13.4	19.2	82^{o}-85^{o}	9.5233
Ib ˌpill	10	0.7	13.2	3.1	10.0	11.9	16.4	82^{o}-85^{o}	8.9726
Ic ˋpill	9	0.7	15.6	3.6	12.0	14.3	20.3	85^{o}-86^{o}	12.7060
Id ˏpill	10	1.1	14.8	3.9	10.9	13.2	18.6	84^{o}-86^{o}	11.1010

4. p-, sp-, b-, -s + b-

4.1. General remarks

Groups II,1; II,2; II,9 were selected for the investigation of the relationships between the above-mentioned sounds (see 1.2.).

Some calculations were made to establish whether the p's and the b's of the 3 groups differ significantly or not.

The results were as follows:

p

		total duration cs	open interval cs	closure cs
II, 1a	peak	17.3	3.9	13.4
		o	*	o
II, 2a	pill	17.8	4.4	13.4
		*	*?	o
II, 9a	purr	18.8	4.8	13.9
		1a–9a: ***	1a–9a: **	1a–9a: o

b

		total duration cs
II, 1c	beak	15.0
		***
II, 2c	bill	13.5
		***
II, 9c	burr	14.9
		1c–9c: o

Accordingly it was decided to make comparisons only between members of sub-groups within the same group (see 3.5.).

4.2. Comparisons in group II,1

a. p-/(s)p-

The segmentation of this cluster presents a problem on account of the co-articulation of the two sounds. However, there is always a clear "hump" in the aerometer curve immediately followed by the sudden drop of the p transition. It was decided to place the dividing line between s and p just after the vertex of this hump. At this point the s noise in the microphone oscillogram seems to have weakened very much or to have stopped altogether.

The difference between the means of the total duration of p (IS) in peak and speak is highly significant (at the 99.9 % level). Statistically, the two stops, which occur in comparable rhythmic-tonetic groups, do not belong to the same population. The difference must be ascribed to phonetic "causes": s precedes p.

The duration of the open interval of (s)p in this group is only about 1/3 of that in p. Besides the intensity of the aspiration is slight, it may even be inaudible. It is, however, in all cases a stable physiological component of the "weakened" p. Also the means of the closure (10.4 cs and 13.4 cs respectively) differ significantly (at the 99.5 % level).

The h values show that, on the average, the relative maximum airflow in (s)p- is only about 2/3 of that in p. Similarly, the relative increase in airflow per unit of time is smaller in the former than in the latter (see table IV, section 4.5 below).

The most important difference, probably, between the two sounds (and very likely the one from which all other differences derive) is found in the glottal gesture: p- has its own gesture with a duration of about 17 cs, whereas s and p in sp-, as it were, share a glottal gesture of somewhat longer duration, 23 - 24 cs. The outstanding feature is that s always occupies the whole of the opening phase and more or less of the closing phase of the glottal cycle, so that the implosion of p begins only during the latter phase.

The plosion (point E) takes place late, in fact only when the closing has been accomplished (the curve of the Photoglottograph has become "horizontal" some cs before E), from which it will be understood that the open interval is short, and the aspiration weak: at the time E the glottis is being adjusted for phonation. This adjustment is achieved when - after the release - the supraglottal pressure has dropped sufficiently. But in E this has not yet taken place. - Kim (1970 p. 114) states that "by the time /p/ is released, the glottis will already have become so narrow that the voicing for the following vowel will immediately start, and then we have an unaspirated /p/ after /s/". If this is to be taken as a general rule, I disagree. My photoglottograms show very clearly that as far as the onset of voicing is concerned, the statement does not hold generally, at least not for my (British English) informant.

Everywhere in group II, lb it has been possible to draw an acceptable and consistent dividing line between the s and p segments.

b. p in sp- and b in -s + b-

It has been demonstrated in section a. above that the characteristics of p when preceded by s are altered radically. Phonetically, the sound comes closer to b in many respects. The question is whether it is identical with b from an articulatory point of view.

The difference between the means of total duration of p in speak (11.8 cs) and of b in Jack's beak (12.8 cs) is significant at the 99.5 % level.

Now the previous investigations have shown that quantity in phonetically comparable groups is greatly influenced by differences in the rhythmic-tonetic structure of the sentences. Unfortunately, at the outset I was unaware of the exact nature of these facts (but had a suspicion of them). It follows that the appropriate test sentences were not made.

However, from a durational point of view (s)p- is neither p nor b, total duration of p-, (s)p-, b and (-s+)b- being 17.3, 11,8, 15,0 and 12,8 cs respectively.

As regards aspiration and voicing, the two sounds also differ from each other: in (s)p- aspiration is a stable feature, whereas in (-s +) b - it is mainly unstable or missing. Moreover, there is a somewhat unstable, but nevertheless unmistakable tendency towards voicing in the EM phase in (-s +)b-. In (s)p-this tendency is entirely lacking during the release stage.

Only in 7 cases out of 9 did the Photoglottograph show a glottal gesture underlying the k-s-b segments (and even these seven were not all very clear).

There seems to be no evidence of any particular artefacts (e.g. that the tongue or the epiglottis may cut off the light from the light source) associated with the photo-glottography of sequences like s + labial + high front vowel (Frøkjær-Jensen et al. 1971). Hence a tentative interpretation of the timing of b relatively to the glottal cycle may be ventured: b only occupies a small part of the glottal gesture, if any, in the articulation of k-s-b. As far as can be ascertained, its implosion falls very late in the closing phase, and no separate gesture underlies the hold and release stages, during which the curve is "horizontal". This feature, too, distinguishes (-s+)b- from (s)p-. The fact that the sounds of the sp- cluster share a glottal gesture, whereas b- is located outside or "at the edge of" the glottal gesture for -k-s might, perhaps, be interpreted to the effect that there is a juncture effecting a less intimate connection between the sounds in -s + b- than in sp-.

The "two-humped" shape of the few tolerably clear photoglottograms of the series II, 1d (Nos. 1, 4, 5, 6) leads one to suppose that -k-s are less intimately connected (are there 2 commands?) than sp-, which only seems to have one command, and that b- has no command at all.

In this sub-group there is no voicing during the hold stage.

c. b in beak and b in Jack's beak

The difference between the means of the total duration of the two b's is very probably significant (exceeding the 95 %, but not the 97.5 % level). This agrees well with previous findings.

d. Total duration of sp- in speak and -s+ b- in Jack's beak

As they occur in separate tonetic groups, and as the rhythmical structure of the sentences is not the same, these two values are not directly comparable. Statistically, the latter (s = 0.8) is a more homogeneous group than the former (s = 1.6). The difference between their means is highly significant (at the 99.5 % level).

4.3. Comparisons in group II, 2

a. p-/(s)p-

Here, too, the difference between the total duration is highly significant (at the 99.9 % level). It follows that the two p's belong to separate statistical populations.

The aspiration in II, 2b is more conspicuous in the microphone oscillogram and makes up a larger proportion of the aspiration in II, 2a than is the case in II, 1b compared with II, 1a.

On the whole, (s)p- in this group is more like a p- than the corresponding sound in II, 1b; thus the values of maximum airflow and relative increase in airflow per unit of time are greater and approximate or exceed the values of b- in II, 2c - in contradistinction to II, 1, in which the values of b- exceed those of (s)p- in either respect.

b. p in sp- and b in -s + b-

A glottal gesture for k-s-b was only visible in a few cases in the photoglottogram (always a minor one); in the remaining cases the glottogram suggested "closed glottis" without voice.

In the aerometer curve for Nos. 6 and 7 (II, 2d) the k-s segments were fused in such a manner that segmentation was rendered impossible. This procedure was, however, very difficult in several items of this series. Another peculiarity was a relatively heavy aspiration sometimes occurring side by side with beginning voice in the release stage. Thus this b with its great h and tan. V values is to some extent suggestive of a p-.

For significance, etc., see 4.2 a.

c. Total duration of sp- in spill and -s + b- in Jack's bill

The difference between the two sets of means is highly significant (99.95 % level).

4.4. Comparisons in group II, 9

a. p-/(s)p-

As in the previous two groups, the difference between the means of total duration is highly significant (at the 99.5 % level).

The photoglottograms for p-, whose total duration is the greatest in the 3 groups, show outlines differing much from the usual, rather regular conic section. Three (Nos. 1, 2, 10) show a marked plateau, several cs in duration, instead of the usual "peak", three (Nos. 7, 8, 9) are distinctly "two-humped", whereas only one (No. 6) is of normal shape.

The two-humped shape is very likely the most interesting, as it is difficult to explain from the assumption of only one command per p. How are the movements maximum glottal opening ⟶ diminishing glottal opening ⟶ maximum glottal opening to be visualized from the assumption of only one command?

It is impossible to take a decision on the question whether the apparatus functioned abnormally (small displacements of the light source?), as the effect on the glottogram, if any, of such displacements is unknown.

The total duration of p- in II, 9a was the greatest of the three groups under consideration. Its relationships in this respect to the other two p's is seen in 4.1.).

The outline of the glottal opening for the sp- cluster is decidedly asymmetric, even more so, perhaps, than in the other two groups, so that the opening phase of the cycle is clearly shorter than the closing phase.

b. p- in sp- and b- in -s + b-

As the b's of II, 9d - alone among the b's of the 3 groups - show a marked tendency towards voicing in the IE phase, they deserve special mention. This applies to Nos. 2, 4, 6, 8, 9, possibly also to 1 and 10. This fact is remarkable as it demonstrates that b preceded by s (in open syllables?) is not necessarily devoiced in its entire length, even during the hold stage, where onset of voice occurs in the latter half of the closure.

This may also provide an answer to the important question: what is the fundamental difference between p and b in English.

On the evidence of the photoglottograms investigated for this article my tentative answer to the question would be like this: p is, as it were, "tied" to a glottal gesture in all positions (with potential aspiration as a consequence), whereas b is "free" in this respect. The supraglottal pressure having risen to the critical level, phonation ceases in b, but the vocal chords do not part, and so phonation is resumed whenever permitted by a sufficient drop in the oral pressure.

This is seen clearly in the case of (s)p- and (-s +)b- in groups 9b and 9c. In the latter, phonation is frequently resumed during the hold stage. This phenomenon is never seen in (s)p-: in this cluster p is also "tied" to a glottal gesture shared with s.

With some confidence I therefore venture to say that a "p" is a [p] also when preceded by s, and a "b" is a [b].

The difference between the means of b and p is not significant.

c. b in burr and b in Jack's burr

The difference between the means of the two sounds is highly significant (exceeding the 99.5 % level).

d. Total duration of sp- in spur and -s + b- in Jack's burr

The difference between the two sets of means is highly significant (exceeding the 99.5 % level).

4.4. Conclusion

In the preceding paragraphs an attempt has been made to find an answer to the question: can (s)p- be looked upon as a b on articulatory grounds, phonetically and phonologically? In other words, is the notation /sbi:k/ "as good as" the notation /spi:k/?

The answer must be in the negative on grounds deducible from the phonetic facts brought to light by this investigation.

Admittedly, the total duration of p is shortened drastically in the sp- cluster, which brings (s)p- nearer to b. But in spite of any such alterations (s)p- potentially retains its aspiration actualized in a measurable quantity in 28 of the 29 specimens investigated. Moreover, (s)p-'s relationship to the underlying glottal gesture is different from the conditions obtaining in both b- and -s + b-. This is most important.

In a similar way b is changed after s (devoicing, tendency towards aspiration in the open interval), but the sound potentially retains voicing, frequently actualized in the open interval and even during the closure (group II, 9 d). Furthermore, it may be supposed that the total length of b is not materially shortened after s. Finally, that unlike p it has no separate command.

But perhaps the question itself is posed the wrong way, and the answer consequently inadequate or inconclusive.

Perhaps one must say that under such and such conditions a given sound will be perceived as p (in English), whereas the identical sound under other conditions would be perceived as b.

This might be called a relational approach. As regards labial stops it implies that what is heard not only depends on what is actually said, i.e., what can be said about the utterance in physiological and acoustic terms, but also on the ears that hear, and on the relational conditions under which the utterance is said and heard (perceived).

And linguistic "relativity", it would seem, not only applies to the total duration of stops, but to any speech sound.

It was demonstrated in 3.5. that the total length of pill in Ic and pill in II, 2a varies significantly with the rhythmic-tonetic conditions obtaining in the sentences. The same is the case with the [I] segment in the same two groups. But even vowels seem to be subject to "relativity". The length of the [I] segment in bitter (group I) can easily be measured in the aerometer curve. The means of the four groups are respectively a: 5.1 cs; b: 6.2; c: 4.9; d: 5.0. The difference between the means of a and b is significant (at the 99 % level), s = 0.7 (for a) and = 1.0 (for b). Other calculations were not made.

Seen from a relational point of view the question posed above concerning the identity of p in the sp cluster actually becomes a pseudo-problem.

If in a given relation (or context), e.g. in the sp cluster, the sound is perceived as a p, it must be accepted that the sound is a p in that context, even if a certain amount of physiological-acoustic evidence could be adduced to the contrary. Which is hardly possible in this case (see above).

Consequently, if such a p is isolated from its relations, by removal of the s noise, e.g., one might very well anticipate b perceptions (in English). From a relational point of view it would seem possible to predict this kind of perception with appreciable accuracy on condition that one knew the rhythmic-tonetic conditions under which the word was spoken. But the amputation of s brings about new contexts, new significant relationships, probably. And what importance can then be attached to the result of such an operation?

If this is so, the ground is cut from under the phonological discussion in favour of the interpretation [sp] ⟶ /sb/.

4.5. Survey of group II

TABLE IV

a) subgroups 1a, 2a, 9a

p in	N	s	total duration in cs	open interval in cs	closure	IF	h	V (slope)	tan. V
\`peak	9	1.3	17.3	3.9	13.4	15.7	22.8	85^{o}-87^{o}	14.7252
\`pill	10	0.8	17.8	4.4	13.4	16.4	24.6	86^{o}-87^{o}	17.1690
\`purr	10	1.0	18.8	4.8	13.9	17.2	25.4	80^{o}-87^{o}	13.3884

b) subgroups 1b, 2b, 9b

sp- in	N	duration of s	total duration of p	s	open interval	closure	EF	h	V (slope)	tan. V	s + p
ˎspeak	9	9.7	11.8	0.6	1.4	10.4	1.5	15.3	83°-87°	11.4893	21.5
ˎspill	10	11.6	12.3	1.2	1.5	10.8	1.6	17.2	85°-87°	14.9699	23.8
ˎspur	10	11.3	13.0	0.9	1.8	11.1	1.9	17.8	83°-86°	11.6756	24.2

Of a total of 29 p's 10 items had relatively strong aspiration, 18 items had weak aspiration, and 1 could not be decided.

c) subgroups 1c, 2c, 9c

b in	N	s	total duration	open interval	closure	h	V (slope)	tan. V	fully voiced
ˎbeak	9	0.9	15.0	1.5	13.5	16.3	84°-87°	13.3433	2
ˎbill	10	0.9	13.5	1.3	12.2	17.8	85°-86°	12.8655	4
ˎburr	10	1.2	14.9	1.4	13.5	18.0	82°-87°	13.5395	6

d) subgroups 1d, 2d, 9d

-s + b in	N	duration of s	total duration of b	s	open interval	closure	s + b	h	V (slope)	tan. V
'Jack's 'beak	9	3.3	12.8	0.7	2.0	10.8	16.1	15.5	83°-87°	11.3827
' - 'bill	8	4.3	12.8	0.7	1.6	11.2	17.1	17.2	85°-87°	14.4194
' - 'burr	10	5.0	13.4	0.5	1.8	11.6	18.4	17.7	80°-87°	11.0429

5. Niels Davidsen-Nielsen: English Stops After Initial /s/ (1969a, 1969b)

As stated in 1.1, one of the impulses to investigate English labial stops came from the above-mentioned article (1969a), whose approach seemed to me to be open to question.

It says (p. 56): "The words were inserted into sentences and recorded by four persons (two English and two American) ... The test words from column I (among them the word spear, H. P.) were then cut out of their environments. In the six words thus isolated the initial [s] was removed ...

The perceptory experiment consisted in letting 32 test persons (24 English and 8 American) identify 52 recordings of these truncated words. Each of these words, which had been randomized, was played twice to the test persons, who were then asked to write down the English word they thought they heard".

The phonetic surroundings of the 6 test words are not identical (1969b p. 324), and it may be questioned whether in all cases the rhythmic-tonetic patterns are the same. This may have had some (perhaps negligible) influence on the total duration of the stops. More important, however, are the implications of cutting experiments as such.

(a) The amputation of s, e.g. in the word spear, interferes with the inner relational conditions of the word and thereby creates new relations, which as likely as not do not exist in identical form in the English language, and which at any rate did not exist in the word before the amputation.

(b) By the removal of s a new initial stop, unknown to the English linguistic system, comes into existence. No matter how natives may perceive that sound - as b or p - this perception says nothing about (s)p-. For the new labial stop is non-existent in English, so to speak. And units being non-existent in a language must be rejected as elements in linguistic reasoning bearing on that language.

(c) If nevertheless the new stop is presented as belonging to the phonetic pattern of English it can be predicted with good certainty that it will be perceived as b since, in the main, no alternative is left open.

In order to substantiate the assertions made above the words purr, spur, burr were selected. They were said in the carrier sentence I say ˋ... today. It will be seen that the 3 words occur in identical surroundings, phonetically, tonetically and rhythmically. Besides, the relations within this group do not differ much from the ones that would obtain in the group pier, spear, beer of Davidsen-Nielsen's experiment if these words had occurred in a similar sentence.

The average total duration of the words purr, spur, burr is 45.6 cs, 45.2 cs, 51.9 cs respectively. If s is cut away from spur, it leaves a "truncated" word with a total duration of 40.6 cs. This word is called t. The relations obtaining in the 4 words are as follows:

$$\frac{p}{purr} = 0.41; \quad \frac{b}{burr} = 0.33; \quad \frac{p}{spur} = 0.25; \quad \frac{p}{t} = 0.32;$$

$$\frac{closure}{purr} = 0.31; \quad \frac{closure}{burr} = 0.30; \quad \frac{closure}{spur} = 0.21; \quad \frac{closure}{t} = 0.27;$$

(see 4.5)

It will be seen

(a) that the amputation of s has created new relations in t which did not exist in spur before the amputation

(b) that a new initial stop has come into existence

(c) that the relations obtaining in t are much nearer to those obtaining in burr than to those obtaining in purr.

As voicing/voicelessness is of minor importance for the differentiation of /b/ - /p/ initially, and as the relations between the durations in the truncated word agree pretty closely with those in burr (but differ widely from those in purr), it is possible - in so far as these relations are decisive for the perception - to predict with a high degree of probability that the number of b perceptions will be great compared with the number of p perceptions.

To this must be added that the open interval of p in spur (1.8 cs) does not differ much from that of b in burr (1.2 cs), but differs widely from that of p in purr (4.8 cs). Even in isolation this fact should greatly increase the probability of b perceptions in the truncated word (Lisker & Abramson 1964).

The conclusion to be drawn from the above and from Davidsen-Nielsen's experiments is that an initial labial stop characterized by relations elsewhere unknown to the linguistic system of the English language, but whose characteristics are similar to those of b in the same position, will be perceived as b with very great probability. For it is a well-known fact that when people are faced with linguistic relations unknown from their mother tongue, they will perceive a given foreign speech sound as "a something" whose relations are nearest to those of a speech sound known from the mother tongue (cp. that no doubt most Danes without phonetic training would spontaneously perceive b in the local pronunciation of the word Palermo; but this is neither a phonetic nor a phonological proof that the initial sound of that place name really is a b)

This, however, says nothing about the natural relations of p in the sp cluster and, consequently, nothing about the phonetic characteristics and phonemic status of (s)p-.

On the whole, amputation experiments seem to me to be methodically inadmissible as evidence in phonological discussions. On the other hand, they are of course quite legitimate in perceptory experiments whose sole object is to find the cues requisite for a given perception.

Acknowledgements

I wish to thank Professor Knud Sørensen, of Aarhus University, who has revised the draft of this article; my son-in-law Niels Glerup Sørensen, who has drawn the figures; the staff of the Institute of Phonetics, especially Mr. Jørgen Rischel, for their kind assistance.

References

Abramson, A.S., L. Lisker and F.S. Cooper 1965: "Laryngeal Activity in Stop Consonants", Speech Research Nov. 1965, p. 6.1.-6.13.

Chomsky, N. & M. Halle 1968: The Sound Pattern of English New York.

Davidsen-Nielsen, N. 1969 a: "English Stops After Initial /s/", Aripuc 3/1968, p. 55-62.

Davidsen-Nielsen, N. 1969 b: "English Stops After Initial /s/", English Studies, Vol L,4, p. 1-19.

Davidsen-Nielsen, N. 1970: Engelsk Fonetik, Copenhagen.

Denes, P. 1955: "Effect of Duration on the Perception of Voicing", JASA, Vol. 27,4., p. 761-764.

Fischer-Jørgensen, E., 1957: "What Can the New Techniques of Acoustic Phonetics Contribute to Linguistics", Proc. VIII Intern. Congress of Ling., Oslo, p. 433-479.

Fischer-Jørgensen, E., 1962: Almen Fonetik, Copenhagen.

Fischer-Jørgensen, E., 1970: Indledning til Experimentalfonetisk Kursus, Copenhagen.

Frøkjær-Jensen, B., C. Ludvigsen and J. Rischel, 1971: "A glottographic study of some Danish consonants", Form & Substance, Copenhagen, p. 123-140.

Gimson, A.C., 1964: *An Introduction to the Pronunciation of English*, London.

Jespersen, O., 1958: *English Phonetics*, Copenhagen.

Jones, D., 1962: *An Outline of English Phonetics*, Cambridge.

Kim, Chin-Wu, 1970: "A Theory of Aspiration", *Phonetica* 21,2.

Kingdon, R., 1966: *The Groundwork of English Intonation*, London.

Lisker, L., 1957: "Closure Duration and the Intervocalic Voiced-Voiceless Distinction in English", *Language* 33, p. 42-49.

Lisker, L. and A.S. Abramson 1964: "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurement", *Word* 20,3, p. 384-422.

O'Connor, J.D. 1963: *A Course of English Intonation*, Stockholm.

O'Connor, J.D. & G.F. Arnold 1966: "*Intonation of Colloquial English*, London.

Thorsen, N.G., 1971: "Voicing in British English *t* and *d* in Contact with other Consonants", *Aripuc 5*, p. 1-39.