WHAT IS THE DIFFERENCE BETWEEN ENGLISH AND SWEDISH DENTAL STOPS?

Sidney Wood

During the past few months Gösta Bruce and I have cooperated in the production of several X-ray motion films of speech movements for investigations of vowel articulations and cavity formation. Each of us has his own informants and his own specific problem area, but the practical details have been arranged jointly. While we have hardly had time to analyse the films in depth it is already clear that they reveal many details of interest for the study of speech production over and above the vowel movements they were designed to portray. One such detail is the laminal articulation of dental stops by an English subject in contrast to the apical articulation of a Swedish subject.

Swedish and English dental stops

The English dental stops are usually described as having an alveolar place of articulation. There is a phonemic contrast of voice, /t – d/ (<u>tin</u> – <u>din</u>), but there is no contrast with other dental stops. The nearest contrasting stops are the affricates /č, ǰ/ (<u>chin</u>, <u>gin</u>), usually described as palato-alveolar. The Swedish dental stops are usually said to be **truly** dental or post-dental. Here too, the only contrast is one of voice, /t – d/ (<u>tal</u> – <u>dal</u>), not of place. In central and northern dialects (where /r/ involves the tongue tip) all dental consonants are retracted after /r/ with which they coalesce. In this respect Swedish (like Norwegian) resembles a number of the languages of India. The retracted set is referred to as supradental or retroflexed. The sandhi character of this process is obvious across word and morpheme boundaries as in <u>har du</u> ([haɖɯ]) but within morphemes there are minimal pairs such as <u>bod</u> – <u>bord</u> ([bu:d] – [bu:ɖ]). It is a matter of linguistic creed whether the minimal pairs are used as evidence for a phonemic contrast between retracted and non-retracted dental stops (/būd/ – /būɖ/) or are generated by the same sandhi rules as are necessary at morpheme boundaries (/būrd/ ⟶ [bu:ɖ], <u>bord</u>).

However, the Swedish subject described in this paper does not represent this type of dialect. He has instead a southern dialect with a dorso-uvular [ʁ] for /r/ and therefore lacks the retracted set of dentals, the sandhi rules not being applicable in this case. Nevertheless, it is still necessary to bear the retracted supradental set in mind during a general discussion of Swedish dentals, and especially for a comparison with English, since many Swedes unfortunately seem to prefer the retracted set when speaking English.

The traditional view expressed in pronunciation handbooks is that the English stops are articulated further back along the tooth ridge. This well established belief ought to be founded on reliable evidence since there is excellent tactile sensation in the tongue tip and it is relatively easy to make palatograms and linguograms.

Swedes are very conscious of an audible difference between English and Swedish dental consonants and look upon the quality of the English stops as a typical ingredient of an English accent in Swedish. Being so keenly aware of such an audible difference, they tend to be concerned not to let it tarnish their English pronunciation. My own personal view is that a Swedish [t] or [d] is more acceptable in English than the English consonants would seem to be in Swedish, but I cannot say whether other Englishmen are equally tolerant. Some Swedish language teachers instruct their pupils to aim at the retracted supradental reflex of /...rt.../ or /...rd.../ as an approximation to the English sounds. I find this gives a distinctly foreign quality to their English speech. Johansson (1967) has taken still radiographs of Swedish dental and supradental consonants, and found the latter to be completely retroflexed, with occlusion made by the underside of the tongue tip. Gimson (1962: p. 159) advises foreign students of English against the use of retroflex stops "since these sound over-retracted to the English ear".

Although both the X-rayed subjects have the same alveolar place of articulation for the dental stops, the Swedish subject still finds the English subject's [t] typically foreign in Swedish. The only apparent articulatory difference is between laminal and apical occlusions. Unfortunately, it is impossible to know how typical these two types are for the respective languages, there being no statistical evidence available to confirm any language-specific preference. It is tempting to speculate whether or not the perceived difference is in some way related to the articulatory difference between apical and laminal occlusions. Acoustically, there may be differences in the formant frequency transitions in adjacent vowels, in the spectral character of the plosive burst and in the duration of the burst and the timing of the voice onset of the following vowel.

Apical and laminal stops in the literature

Jones (1956: § 515) described the laminal articulation as a less common variant of the purely apico-dental, providing "a very unnatural effect" in the English speech of "many foreign people, eg French, Italians, Hungarians and some Germans". He describes the English stops as apico-alveolar. Jespersen (1899: §191) records an observation by Passy that French dental stops tended to be laminal. Jespersen recognized that the laminal articulation occurred frequently as an individual variant but he doubted whether the audible difference could be so great that a given language would prefer the one or the other. Malmberg (1967, pp. 95-96) describes the laminal articulation as an individual variant or as dialectal (frequent in central and northern dialects of Swedish and in French). He considered the acoustical differences to be insignificant. Catford (1968: p. 327) lists lamino-dental to lamino-post-alveolar stops as "variants of apical stops occurring in English and elsewhere". The traditional view is thus against a language-specific preference for either articulation apart from the apical being normal and the laminal a variant and against any audible difference between the two.

On the other hand, Ladefoged (1954: p. 19) has found phonemic contrasts between apical and laminal stops in the West African languages Twi, Ewe, Temne and Isoko, which indicates that it must be possible to obtain a useful acoustic difference in this manner and that some languages do categorize the two articulations. Chomsky and Halle (1968: §7.4.4) generalized from this and similar evidence and proposed a contrast between sounds with long constrictions (denoted [+distributed]) and sounds with short constrictions (denoted [-distributed]). According to their generalization, two contrasting places of articulation within the denti-alveolar region are always associated with a difference in the length of the constriction. A new case pointing in the same direction has recently been reported by Nihalani (1974) who has made a palatographic and radiographic examination of the stops of Sindhi. He disagreed with the standard view that the retracted dentals of Sindhi are retroflexed

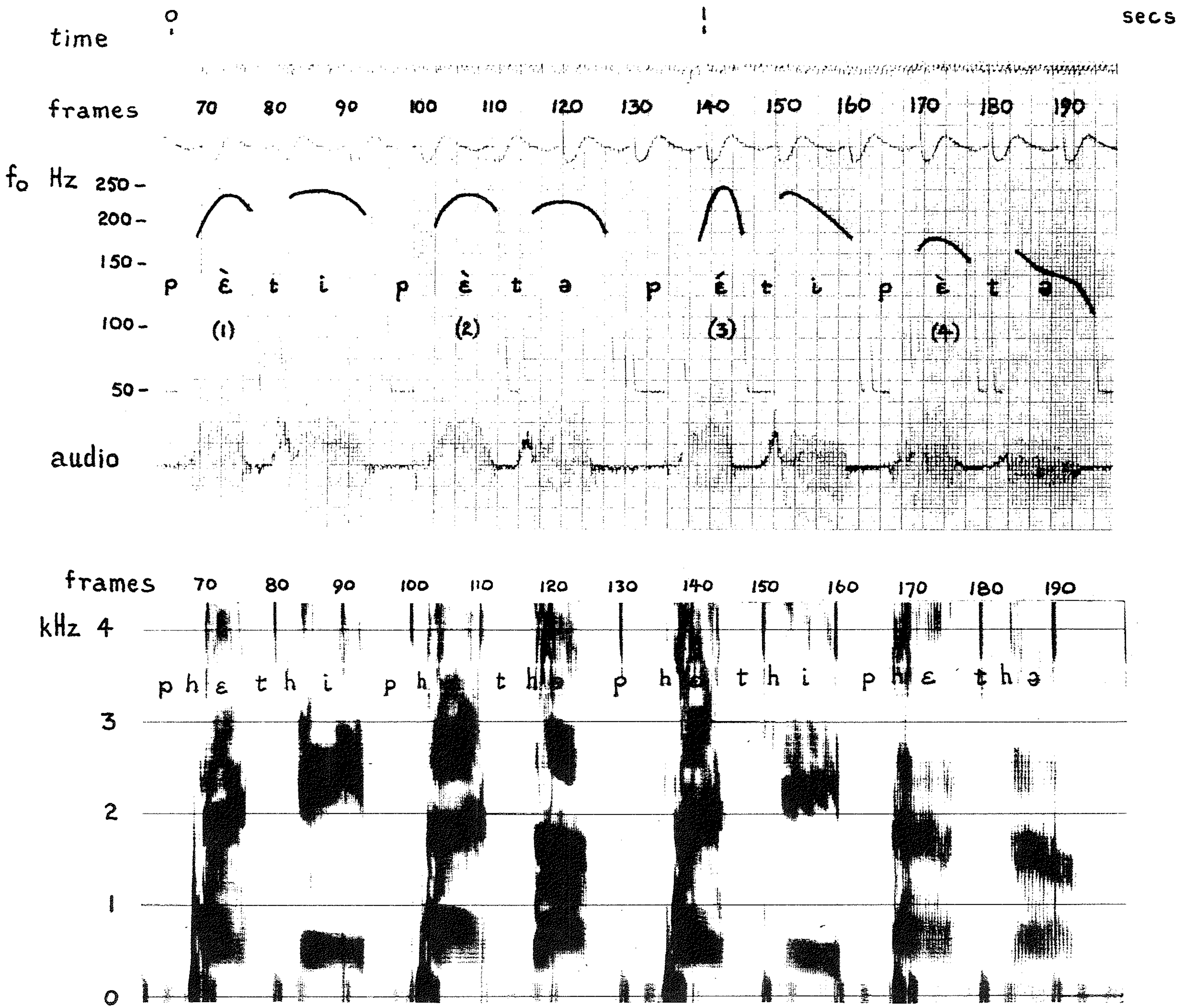

Fig. 1. (a) Oscillograms and (b) spectrograms of one of the English sentences, illustrating the synchronization pulse that appears at every tenth frame on the film.

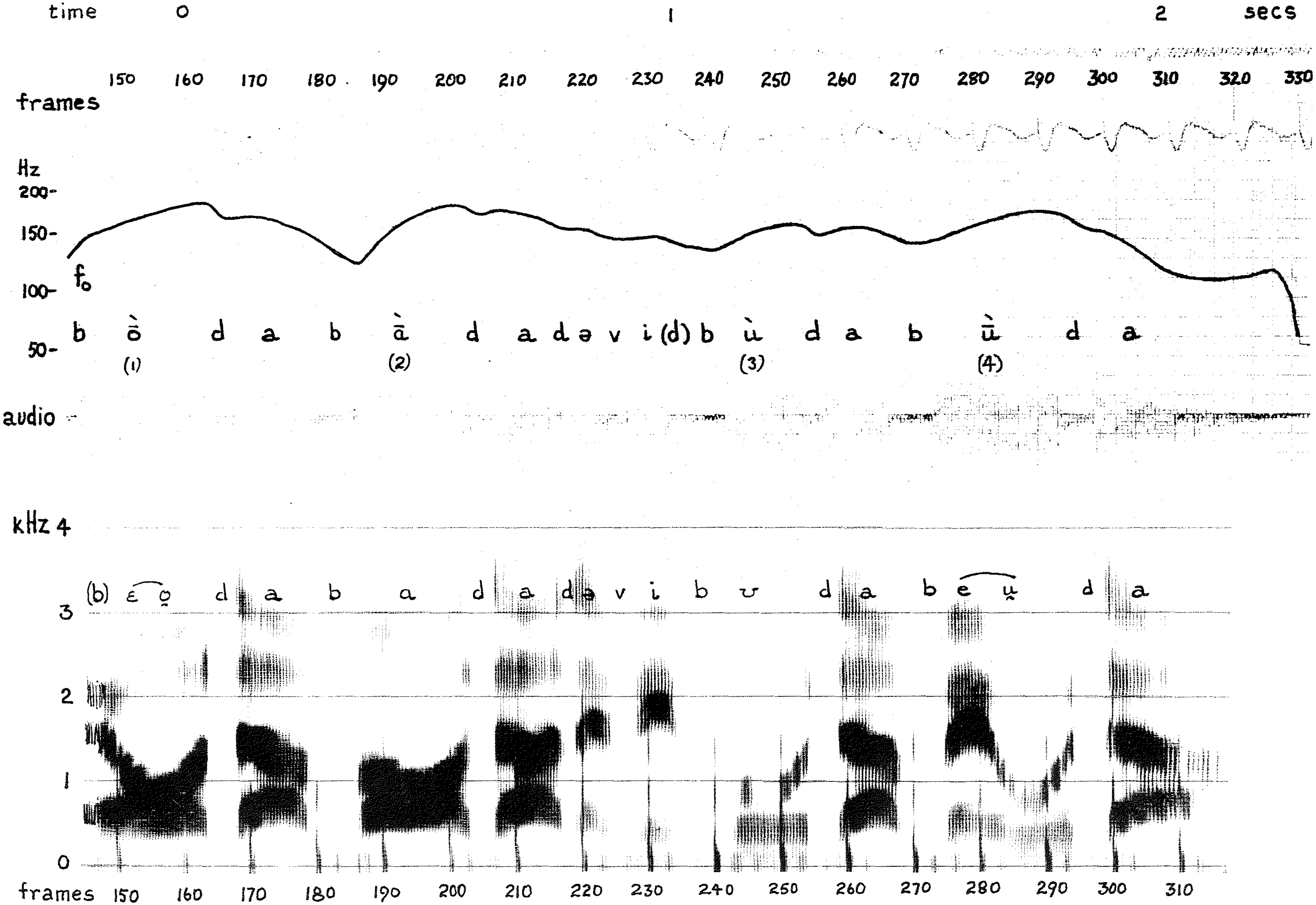

Fig. 2. (a) Oscillograms and (b) spectrograms of one of the Swedish sentences, illustrating the synchronization pulse that appears at every tenth frame of the film.

as in other Indian languages. He found that the non-retracted stops [t, th, d, dh] were laminal with a denti-alveolar location, while the retracted set were apical and post-alveolar but not retroflexed.

Cineradiography

The films were made at the County Hospital, Lund, with the consent of Professor Olof Norman and the assistance of members of the Roentgen Technology Section (Dr Thure Holm, Radiophysicists Gunnila Holje and Gudmund Swahn and Technician Rolf Schöner). The customary high voltage technique was used (120 kV) in order to take advantage of the greater transparency of bone to higher frequency X-radiation, thus producing adequate pictures of both bone and soft tissue.

The laboratory used is primarily intended for angiographic examinations with equipment specialized for observing events in soft tissue such as blood vessels. An added practical advantage is that the camera there provides a synchronizing pulse which appears on every tenth frame and which for clinical purposes can be recorded alongside an elctrocardiogram. We recorded this latter signal simultaneously with the microphone signal, separately on a parallel track. The microphone and synchronizing signals are subsequently processed separately for oscillographic records (Figs. 1a, 2a) but are mixed for spectrograms (Figs. 1b, 2b). We have no opportunity to screen the subject's speech from the relatively high ambient noise level in the laboratory (the high pitch whine of the rotating anode and the rapid intermittent tapping of the camera mechanism). To reduce the recorded level of the background noise relative to the speech we place a highly directional microphone close to the subject's mouth. The noise is audible on the speech recording but barely appears on spectrograms. In order to facilitate enlargement of the radiographs to life size, a long stiff coppar wire with blobs of solder at 1 cm intervals is taped to the subject's forehead and nose and bent so as not to interfere with the movement of the lips and mandible. The camera speed is 75 frames/second (one frame every 13.33 ms) and the duration of radiation is 3 ms per frame. We have 30 to 40 seconds effective time available per reel of film. As a safety precaution each subject is limited to one reel.

The speech utterances that were X-rayed were primarily designed for an investigation of vowel articulation. The present paper is based on articulatory data provided by the consonants framing the vowels. Dental stops (English [t] and Swedish [d]) were repeated four times in each sentence, always intervocalic and following a strong vowel. The utterances conformed to the grammatical structure of a model sentence in which vowels were substituted systematically. Each sentence contained four strong positions in which vowels were substituted systematically. The English sentences were "petty pet a petty-petter, potty pot a potty-potter, ... etc.". For Swedish sets of three vowels were rotated through four strong positions, thus: "Båda badade vid Bodda boda, Boda bådade vid Badda bada, Bada bodade vid Bådda båda, ... etc.". The English sentences were uttered at about 4.5 syllables/second, the Swedish sentences at about 4 syllables/second. A further set of English sentences, uttered at about 6.5 syllables/second, has not yet been analysed and could not be included in the present comparison of dental stops.

This paper is devoted to a comparison of two subjects only - myself (Southern British English, NE Kent) and Gösta Bruce (Helsingborg, NW Skåne).

The dental stops of the X-rayed subjects

A striking difference between the two subjects was in the formation of dental stops. Fig. 3 shows how the English speaker made his dental stop with a flat area of the tongue behind the tip (unbroken line, depicting the moment of occlusion). Fig. 4 shows how the Swedish subject made his dental stop with the very tip of the tongue (unbroken line). These tracings distinguish between mandibular and lingual gestures.

The actual shape of the tongue body at the instant of occlusion depends on the preceding vowel. The broken lines at Figs. 3 and 4 show the preceding vowel positions. These represent the turning point (lasting from 1 to 3 frames or about 10 - 50 ms) between the end of the CV movements and the beginning of the [t] or [d] movements. At these speaking rates (4 - 5 syllables/second, a little more leisurely than everyday speech) the gestures of the dental stops (mandible elevation, tongue tip raising) were not initiated until the vowel configurations had been completely formed. The cases illustrated are for a preceding palatal vowel [ɛ] and two pharyngeal vowels [o] and [ɑ]. For the Swedish subject, [o] is the final portion of a diphthong [ɛ͡o̯] for /ō/ (cf.

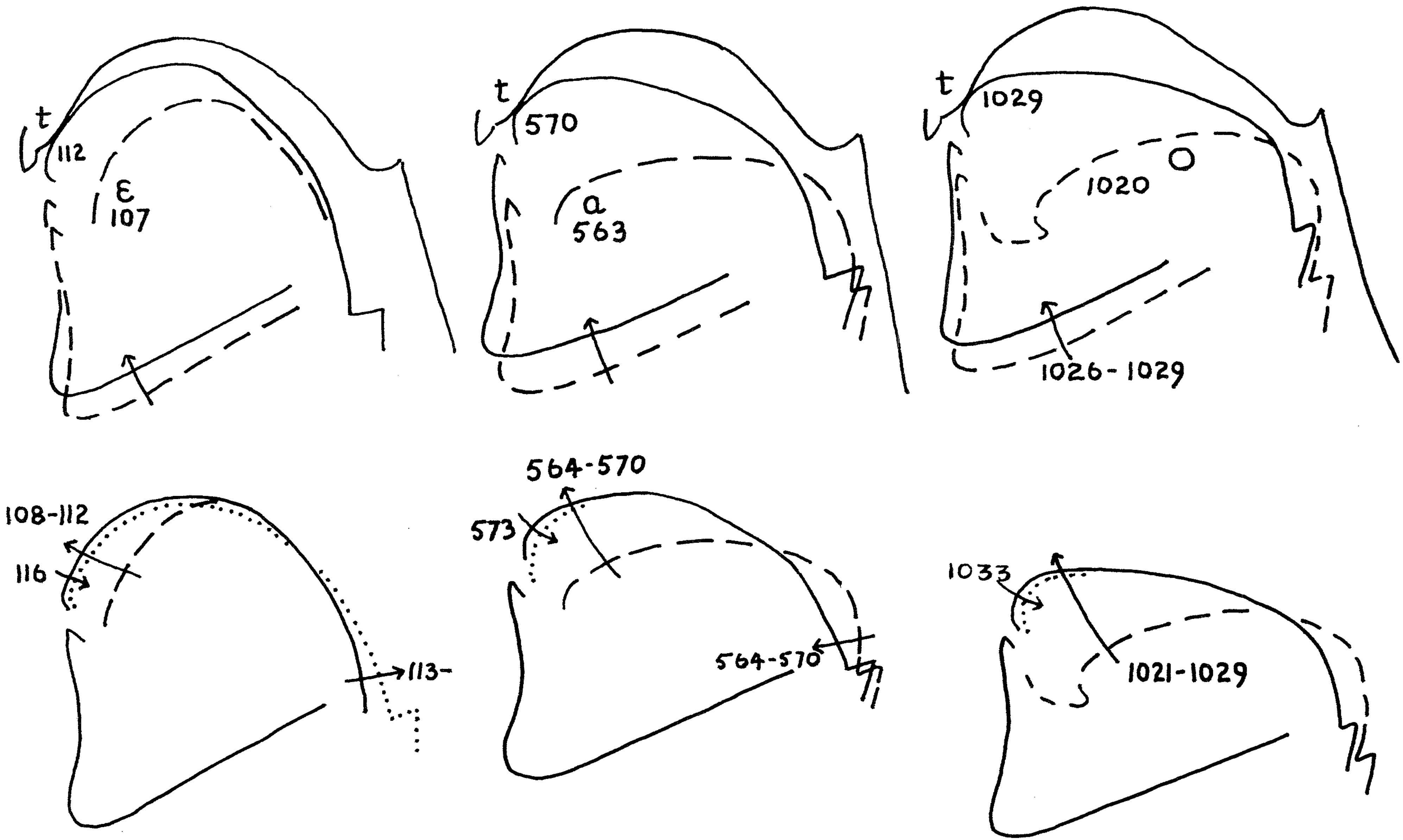

Fig. 3. Tracings of three laminal [t] occlusions after [ε, ɑ, o] by the English subject. Each frame number represents a 13.3 ms interval. Arrows indicate direction of movement.

--- preceding vowel —— occlusion ····· burst

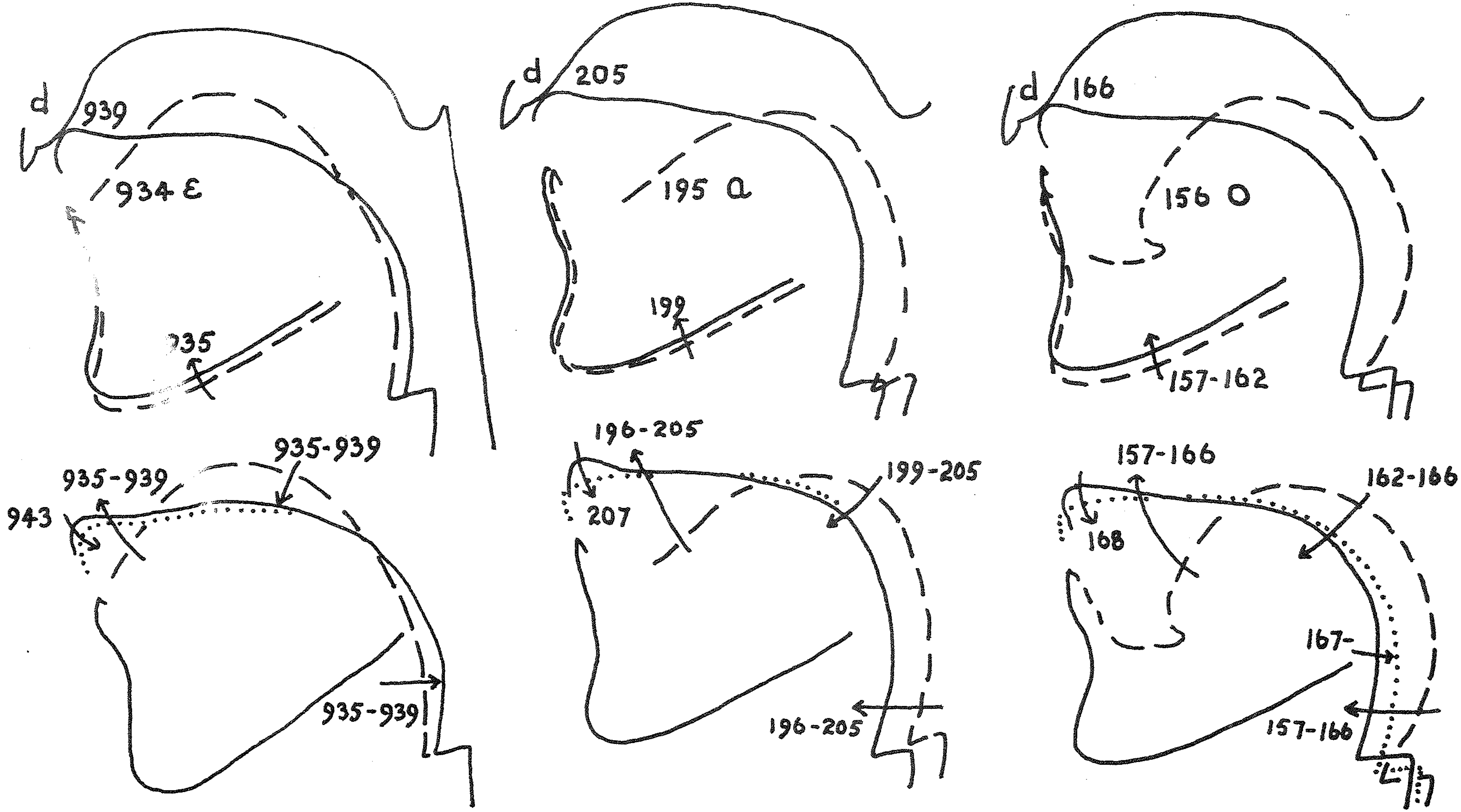

Fig. 4. Tracings of three apical [d] occlusions after [ε, ɑ, o] by the Swedish subject. Each frame number represents a 13.3 ms interval. Arrows indicate direction of movement.

--- preceding vowel —— occlusion ····· burst

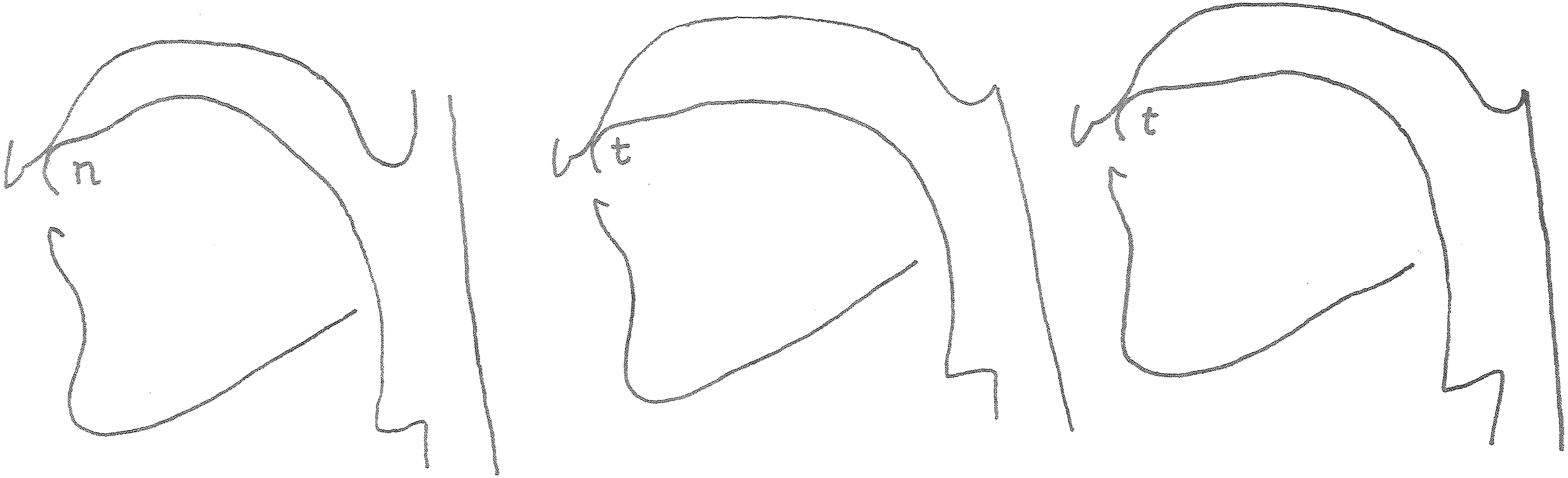

Fig. 5. Tracings of alveolar occlusions by the Swedish subject: [n] in nitton, [t] in nitton and [t] in sjuttio.

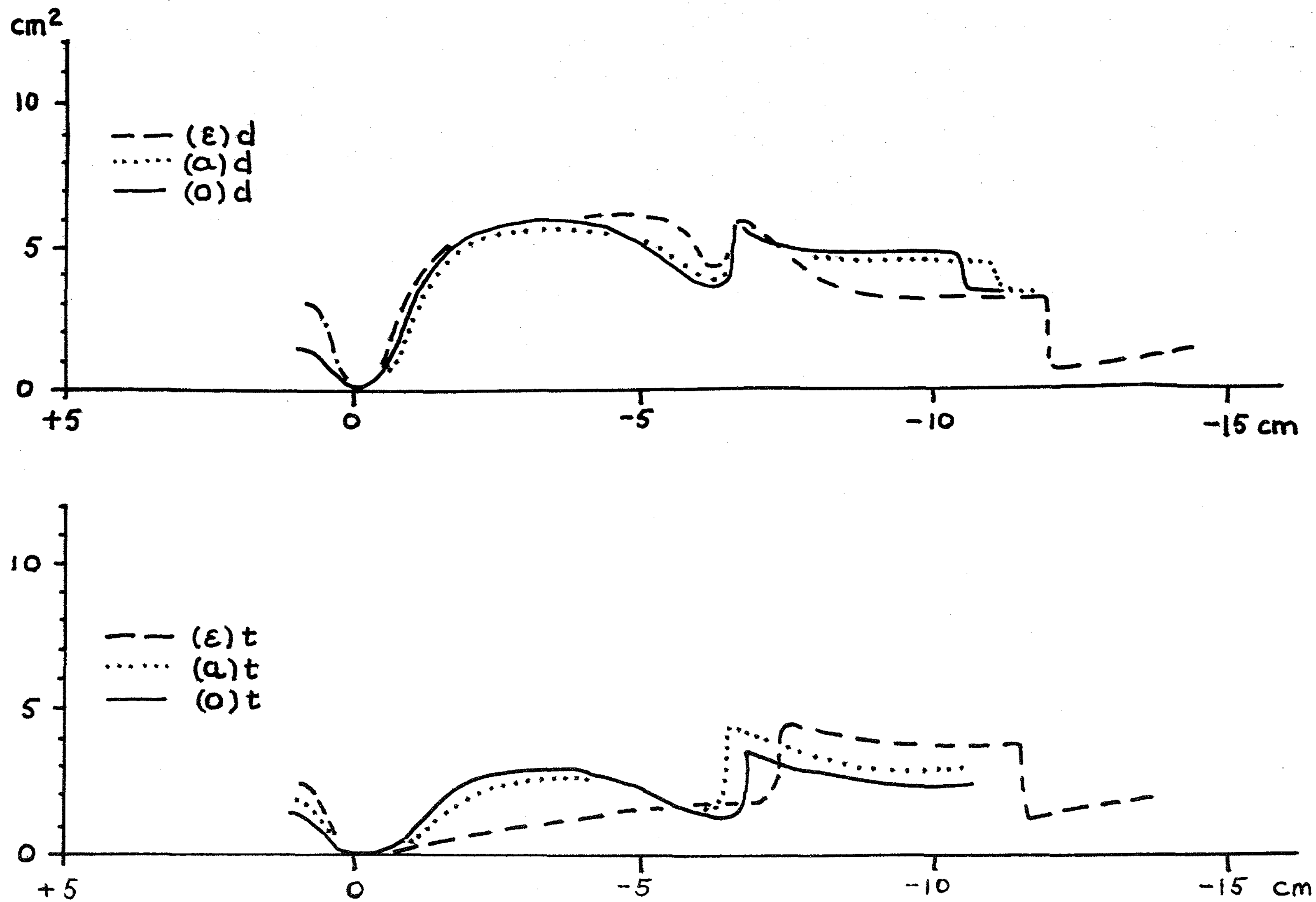

Fig. 6. Area functions of the Swedish subject's apical occlusions at Fig. 4 (above) and of the English subject's laminal occlusions at Fig. 3 (below).

Bruce 1970 and see the spectrograms at Figs. 2b and 7b). Similarly, the shape of the tongue body at the moment of release (dotted lines) depends on the configuration of the oncoming vowel towards which the tongue is moving. To make the examples as comparable as the material will permit, the cases illustrated are [ə] following the English [t] and weak [a] following the Swedish [d].

Comparison of Figs. 3 and 4 indicates practically the same place of occlusion for both subjects. This need not necessarily contradict the general belief in different places of articulation for English and Swedish stops since the Swedish subject is from Helsingborg and Malmberg (1967: pp. 94-95) has reported that the northern part of the province of Skåne is one of several Swedish areas where the dialects have stops that are alveolar rather than dental.

There is also a difference of voicing between the utterances of the subjects ([t] for the English, [d] for the Swedish). Malmberg (p. 95) also stated that there are frequent instances of different places of articulation for voiceless and voiced dental stops, [d] often being more retracted than [t]. He specifically mentioned the south of Sweden generally, reporting that [t] is dental but [d] post-dental. The Swedish subject included his name and the date in his film, and has thus provided examples of other dental consonants. Fig. 5 shows the initial [n] in <u>nitton</u>, the [t] in <u>nitton</u> and the [t] in <u>sjuttio</u>. These are all clearly alveolar and indicate that this subject does not have different places of articulation for [t] and [d].

There is an inherent difference of tongue body shape between the laminal and apical types of articulation. The crown of the tongue arch is clearly higher during the period of occlusion for the laminal articulation than for the apical. This is because the tongue curves upwards behind the laminal occlusion whereas the tongue tip approaches the tooth ridge from below for the apical occlusion. Consequently the vocal tract widens abruptly behind the apical occlusion but gradually behind the laminal occlusion (Fig. 6). The different vocal tract resonance conditions resulting from this will be discussed further on.

In order to compare the acoustical characteristics of [t] segments for both subjects, two separate series of teswords have been specially recorded: ['kvVta] (Swedish) and ['kVtə] (English), where [V] represents a range of vowels for commutation. Most resulting testwords were nonsense

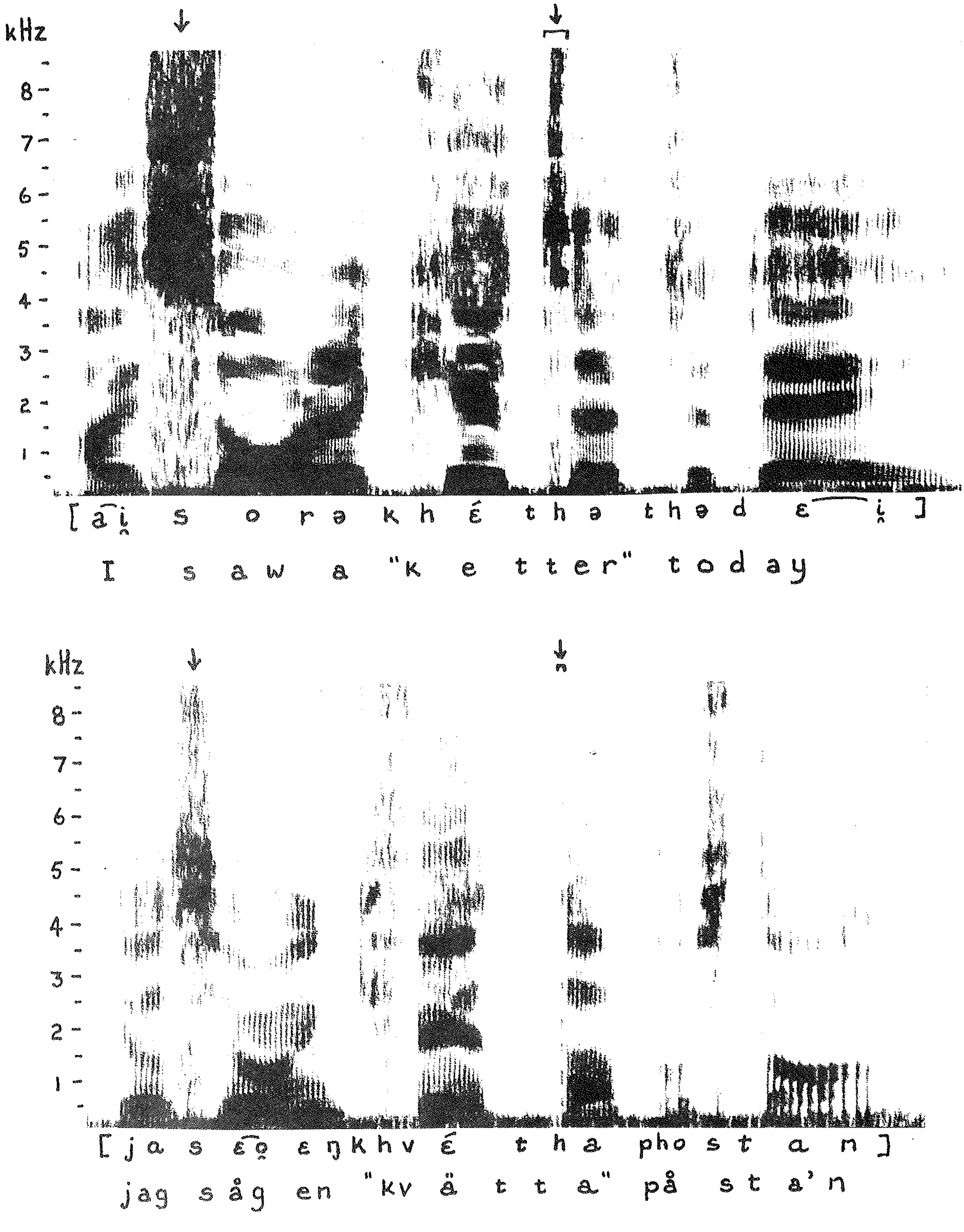

Fig. 7. Spectrograms of two examples of the sentences recorded for a comparison of [t] segments in (a) English (above) and (b) Swedish (below). The arrows indicate the locations of spectral sections (Fig. 8).

but a few were common real words – Swedish kvitta ("settle a debt") and English cutter, carter, cotter (and with a little imagination even kitter, catter, cooter). The velar stop was introduced in order to have a different lingual consonant alternate with the [t]. The fricative [v] was included in the Swedish testword in order to block spirantization of the palato-velar [k] to [ç] before palatal vowels. The testwords were placed in a frame sentence jag såg en ... på sta'n or I saw a ... today, where the [s] segments were also used for comparison.

Spectrograms of one pair of renderings are given at Fig. 7. One very striking difference is the duration of the [t] occlusions in the testwords – the Swedish occlusion was twice as long as the English although the total sentence duration is virtually the same for both speakers. This is related to the characteristically longer postvocalic consonant following short Swedish vowels.

Equally striking are the durations of the [t] bursts – about 40 ms in the English rendering against 10 or 20 ms in the Swedish. There is a corresponding difference in the voice onset. The spectral character of the English [t] burst is similar to that of the [s] – high levels above 4000 Hz with several peaks from 4300 to 7000 Hz. The Swedish burst appears to be weaker and more evenly spread up to about 7000 Hz with negligible levels above while the [s] has two peaks at 4000 – 5000 Hz, somewhat lower levels above and even less below.

There is also a difference in the transitional phases of the vowels towards the [t] occlusion – F2 in the English [ɛ] at Fig. 7 points to a locus at about 1800 or 1900 Hz while it points down more sharply in the Swedish [ɛ] to a locus at about 1600 Hz.

The spectral differences in the [t] bursts can be studied in greater detail at Fig. 8 which shows the spectra sampled at approximately 4 ms intervals (the time scale is very rough). The [s] spectra are included for comparison. This confirms the [s] like character of the English subject's [t] burst, with high levels from 4000 Hz to beyond 8000 Hz. The Swedish subject's [t] burst had a much lower overall intensity, mostly at 4500 Hz and below. The Swedish [t] burst does not resemble the same speaker's [s]. At the same time, the Swedish [s] is different from the English subject's – the Swedish [s] is almost symmetrical around two strong peaks at 4500 – 5000 Hz.

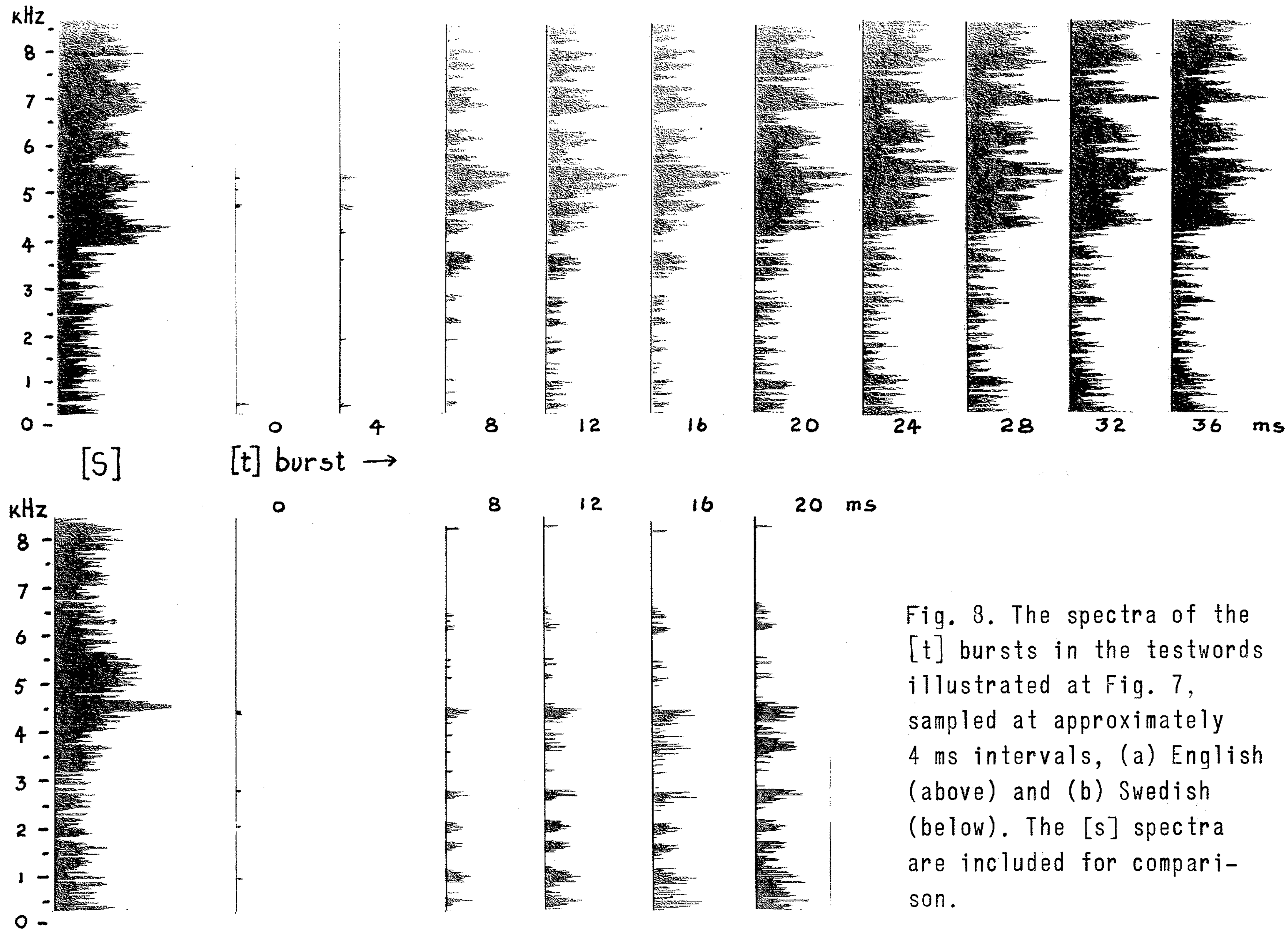

Fig. 8. The spectra of the [t] bursts in the testwords illustrated at Fig. 7, sampled at approximately 4 ms intervals, (a) English (above) and (b) Swedish (below). The [s] spectra are included for comparison.

Discussion

The difference traditionally described is that the English dental stops are more retracted than the Swedish, i.e. there is a difference in the place of articulation. Yet both of the subjects had alveolar stops. Paradoxically, the Swedish speaker had a lower locus, which traditionally would be ascribed to a more retracted place of articulation.

The most obvious difference seen on the X-ray films is that between laminal and apical occlusions, i.e. a difference of tongue shape rather than a difference between the points of aim of lingual gestures. The different vocal tract configurations associated with the two types of occlusion are illustrated at Fig. 6 while the consequent resonance conditions will be discussed below, where it will be argued that the observed loci and spectral differences are compatible with these two types.

There is finally the longer duration and greater overall intensity of the English subject's [t] burst. A necessary articulatory correlate of this is an open glottis to permit an adequate airflow, and a necessary consequence is the delay in voicing onset for the vowel. Glottal adjustment alone would account for a longer aspiration phase and delayed voice. But if the [s]-like quality of the burst is an essential component of the English [t], then the glottal adjustment must also be coordinated with a slow release of the occlusion to maintain a turbulent air stream through the tongue constriction and past the incisors.

Any cross-language comparison based on only two subjects is beset by the problem of how to generalize the results to embrace entire communities, especially when those results are novel and seemingly in contradiction to a long established and widely accepted view. The articulations X-rayed here are in accordance with the belief that English and NW Skåne dental stops have an alveolar place of articulation. Johansson's central Swedish subjects had purely dental occlusions (or more precisely denti-gingival). This is in accordance with the established belief that English and Central Swedish dental stops have different places of articulation. The novelty of the present results is that the dental stops of the two subjects are acoustically different, notwithstanding the same place of articulation. This suggests that the essential difference between English and Swedish dental stops must be something else and not the place of articulation. Gimson's advice to foreign learners of English (pp. 153, 154, 159) points in the same direction - he recommends practising affrication of English /t, d/ as [t^s, d^z]. This advice coincides with the difference observed above between the two subjects - the duration and spectral character of the burst (Fig. 8).

Are the observed acoustic differences related to the difference between apical and laminal occlusions? The traditional view has been that there is hardly any audible difference between the two types of articulation. But the contrasts reported by Ladefoged for certain West African languages indicate that there is a consistent difference that is large enough to be useful for phonemic oppositions. Stevens (1973) has recognized that the character of the turbulence of dental consonants is more dependent on the type of constriction (apical or laminal) than on the location of the constriction. For constrictions in the denti-alveolar region, turbulence excites mainly the front cavity resonances. But his calculations also indicate that a more gradual widening or tapering of the vocal tract behind the constriction results in greater coupling of the turbulence source near the constriction to the back cavity resonances. Consequently, he argues, for laminal articulations several resonances are excited over a range of frequencies above a certain critical frequency determined by the amount of tapering.

Chomsky and Halle believed that their contrast of constriction length would also cover the difference between hard and soft consonants in Slavic languages. They quoted evidence from Polish. Fant has published radiographs and area functions of Russian palatalized (soft) and non-palatalized (hard) consonants, including dental stops (1960: Figs. 2.6-8 and 2.6-9). The articulatory difference made by his subject is precisely that difference between the laminal and apical occlusions described above, and between Chomsky and Halle's long and short constrictions - for the soft dental stops the occlusion was laminal and long with the tongue body high against the hard palate, but for the hard dental stops the occlusion was apical and short with the tongue body depressed.

Stevens's argument is completely in accordance with the data presented by Fant for the Russian soft (laminal) voiceless dental stop relative to the hard (apical) voiceless dental stop (Fant's Fig. 2.6-10). The release burst of the soft laminal voiceless stop was characterized by low spectral levels below 3000 Hz, with a weak F2 at about 1900 Hz, and a high level in the region 6000 - 9000 Hz, while that of the hard apical voiceless dental stop was characterized by a relatively high spectral level below 500 Hz, a zero at 850 Hz, an F2 peak at 1600 Hz and a generally lower level at higher frequencies, declining after smaller peaks in the 3500 - 4500 Hz region.

The similarity between the two Russian [t] types and the [t] bursts of the present subjects (Fig. 8) is very striking. Like the Russian soft [t] burst, the English subject's [t] burst is relatively weak below about 4000 Hz but contains numerous high level peaks throughout the 4000 - 8500 Hz region. This corresponds to the greater excitation of the back cavity resonances when the vocal tract widens gradually behind the constriction, as pointed out by Stevens. Similarly, like the Russian hard [t] burst, the Swedish subject's [t] burst is relatively weak above 5000 Hz. They have peaks at similar frequencies - about 500, 1100, 1600, 3500, 4500 Hz. Unlike the Russian hard [t], the Swedish [t] did not have its maximum level at F2, and it had additional small peaks at 2100 and 2600 Hz. This type of spectrum corresponds to the excitation of the front cavity resonances only and the exclusion of back cavity resonances when the vocal tract widens abruptly behind the constriction, as pointed out by Stevens. The vowel formant transition loci of the English and Swedish [t] occlusions correspond to the F2 frequencies found by Fant in the Russian soft and hard [t] bursts, 1900 and 1600 Hz. The different spectral characters of the [t] bursts of the English and Swedish subjects, and the difference in locus frequency, are thus related to the apical and laminal manners of forming dental occlusions.

Stevens also notes a tendency for the longer laminal constriction to be released more slowly. In the case of voiceless stops the burst is prolonged and the voice onset for the following vowel delayed (he quotes data showing 20 ms or more following the laminal occlusion compared with a few ms after the apical occlusion). This was also observed by Ladefoged in the West African languages, where the laminal stops tended to be affricated. The same difference can also be seen on the spectrograms published by Fant for the Russian consonants (his Fig. A. 13-16) which show a much longer burst after the laminal soft [t], than after the apical hard [t]. The [t] burst of the English subject was nearly three times as long as the Swedish subject's. This difference is thus also in accordance with the laminal and apical types of occlusion.

Conclusions

One firm conclusion at least can be drawn from the present results - that the observed vowel formant transitions, the spectral characters of the bursts and the relative durations of the bursts can be related to the apical and laminal types of occlusion. Would it be presumptuous to assume that Swedish dental stops are generally apical and English dental stops generally laminal? This would contradict the widely held belief that dental stops are apical for most speakers of both languages, laminal for some in both languages.

If the widely held belief is correct, that dental stops in both languages are generally apical, then the difference between laminal and apical articulations observed in this pair of subjects will not be a typical difference between English and Swedish. Nor will the associated spectral differences be typical either and the spectra illustrated at Fig. 8 will then in part be individual speaker-dependent characteristics rathen than essential language-dependent features. What would this leave us with? Firstly, that the English [t] burst is [s]-like whereas the Swedish burst is not. This difference can be achieved with either type of articulation. Secondly, the English burst is longer and voice onset delayed for the next vowel. While this is a typical attribute of the laminal articulation, the latter is not necessary condition for affrication. Gimson's advice to learners, to make the burst [s]-like and affricated, favours both of these differences between English and Swedish. The longer burst requires the vocal folds to remain open for a longer period for English [t], the more intense [s]-like quality of the burst requires greater pressure behind the constriction, while the maintenance of the [s]-like quality thoroughout the whole of the longer burst requires the occlusion to be released more slowly.

How important is the often-quoted difference in place of articulation? For most Swedish dialects the dental stops are said to be truly dental while the English stops are alveolar. The fact that some Swedish dialects have alveolar stops without the "foreignness" of the English stops suggests that the difference in location is unimportant. Stevens has demonstrated that there are regions of the vocal tract where variations of constriction location yield minimal acoustical differences. The denti-alveolar part of the tract is one such region. There are other regions where differences of constriction location yield large acoustical differences, for example from the alveolar to the palatal. Consequently the acoustical difference between alveolar and purely dental consonants is very small, insufficient for a phonemic contrast. In the languages that have contrasting sets of dental stops with different locations, there is always the difference of constriction length (usually apical vs laminal) with the extra redundancy of affrication vs non-affrication. I am sure in my own mind that it is unwise for language teachers to emphasize the retracted English articulation. The danger is that this prompts the pupils to adopt something like the Swedish supradental (retroflexed) consonants. The first priority (eg for French learners) should be to learn aspiration. The next priority (eg for Swedish learners) to make the aspiration [s]-like and prolonged.

References

Bruce G (1970). Diphthongization in the Malmö dialect. Working Papers, Phonetics Laboratory, Lund University, 3:1-19

Catford J C (1968). The articulatory possibilities of man. Malmberg B (ed), Manual of phonetics, pp. 309-333. The Hague

Chomsky N and Halle M (1968). The sound pattern of English. New York

Fant C G M (1960). The acoustic theory of speech production. The Hague

Gimson A (1962). An introduction to the pronunciation of English. London

Jespersen O (1899). Fonetik, en systematisk fremstilling af laeren om sproglyd. Copenhagen

Johansson K (1967). Några perceptoriska försök med svenska dentaler och supradentaler. Unpublished licentiate thesis, Department of Linguistics, Lund

Jones D (1956). An outline of English phonetics. Cambridge. 8th edn

Ladefoged P (1964). A phonetic study of West African languages. Cambridge

Malmberg B (1967). Lärobok i fonetik. Lund

Nihalani P (1974). Lingual articulation of stops in Sindhi. Phonetica 30: 197-212

Stevens K N (1973). Further theoretical and experimental bases for quantal places of articulation for consonants. MIT Research Laboratory of Electronics, QPR 108:247-252