AERODYNAMIC AND ACOUSTIC PATTERNS OF SPEECH

CELIA SCULLY

The final, aerodynamic stage of speech production provides a good place to look for relationships between articulatory and acoustic patterns in a search for the more basic encoding of the speech message. The aerodynamic variable of area A extracted from the raw data of pressure and air flow indicates the timing and degree of closure of the tongue, A_c, and velopharyngeal port, A_n, and gives a rough estimate of the glottal area, A_g, also.

The relationships between articulatory, aerodynamic and acoustic features are likely to prove quite complex. Firstly, for example, a constant supra-glottal feature and a constant sub-glottal feature combined with two different laryngeal features will give rise to two different aerodynamic situations. This may be seen simply by applying Ohm's law to the acoustic circuit of the vocal tract. Secondly, slight but important differences of articulation and hence of aerodynamic conditions can distinguish the production of two phonemes such as /t/ and /s/ in similar phonetic contexts. Thirdly, different articulator movements may be needed in different phoetic contexts to achieve an acoustic feature needed for the perception of a particular phoneme. An example will be cited for each of these three points. The equipment was an electro-aerometer to measure air flow in and out of nose and mouth separately and a tube to measure air pressure in the oral cavity behind a constriction of the vocal tract. The set-up has been described fully elsewhere (Scully 1969).

1

When /s/ and /z/ were compared for one English speaker /s/ had significantly higher oral pressure and air flow but the tongue constriction areas A_c were the same for /s/ and /z/ as were the timings of the tongue movements. So the supra-glottal articulations were not distinguished by a fortis-lenis feature. Yet the open glottis configuration of /s/ gave acoustically fortis features of intensity and duration while the closed glottis of /z/ resulted in acoustically lenis features (Scully 1971). It is wrong to assume that a simple one-to-one transformation always occurs between the articulatory and the acoustic stages of a particular feature.

2.

Some English alveolar consonants have been investigated in detail for a few speakers. Figure 1 shows that the critical articulatory features which distinguish /s/ from /t/ may not be always the expected ones. Figure 1 shows part of two utterances containing /s/ and /t/. Although the tongue constriction area is smaller for /t/ than for /s/, complete tongue closure at the alveolar ridge is NOT indicated during /t/. Also,

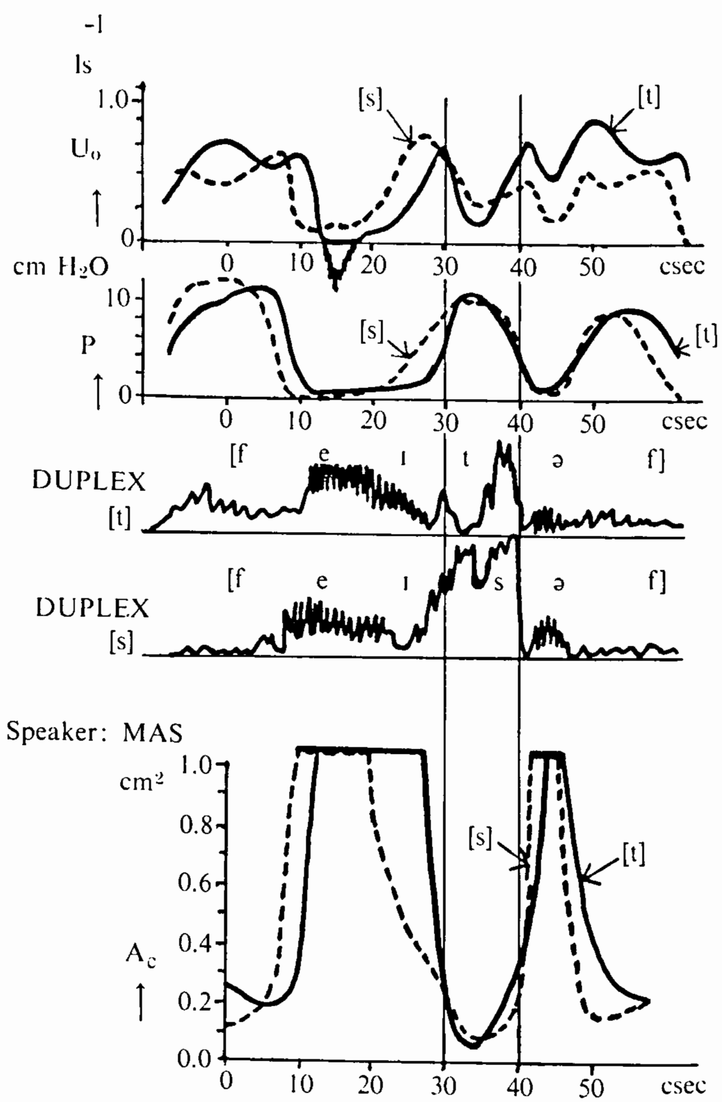

Fig. 1. Oral air flow U_o, oral pressure P, duplex oscillograms (showing intensity of high frequencies) and tongue constriction cross-section area A_c for speaker MAS in 'Say the fate/face afar'. Time markers are 0.1 sec apart.

turbulence noise occurs, not only in the release portion of the /t/ where it is expected, but also during the closure phase. Noise is generated when fairly high air flow and pressure combine. To avoid unwanted noise and hence confusion with /s/, the tongue closure for /t/ must be made rapidly, although the release may be slow and the closure need be only NEARLY complete.

3.

English utterances containing /t/, /d/, /n/, /nt/ and /nd/ in intervocalic position were analysed acoustically for 11 speakers and aerodynamically for a few of these. The time taken to raise the velum for the different speakers varied between about 8 and 18 csec (see Björk 1961). For /n/ alone, as for /d/ alone, the speed of the tongue movement is the rate-determining-factor for the occlusive duration; the velum need not be raised before the tongue release. In the clusters /nt/ and /nd/, however, the velum must be raised DURING the occlusive phase in order to raise the oral pressure sufficiently for a /t/ or /d/ burst to be perceived. Then the velum, not the tongue, is the rate-determining-factor. As a result, /n/ and /d/ are often expanded, not compressed, when in a cluster; this was true for nine out of the eleven speakers. Figure 2 illustrates this: In 'feign', the velum is not raised until the following /f/ segment,

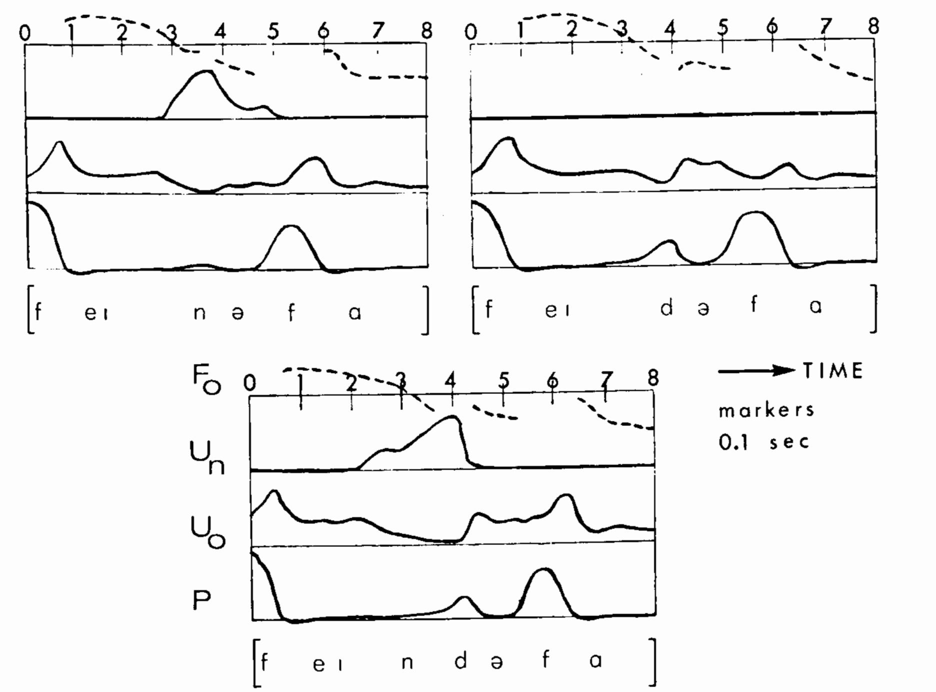

Fig. 2. Fundamental frequency F_o, nasal air flow U_n, oral air flow U_o and oral pressure P for speaker ELT in "Say the feign/fade/feigned afar". Time markers are 0.1 sec apart.

which requires it, whereas in 'feigned' the velum is up at the /d/ tongue release. /n/, /t/ and /nt/ behave similarly, but the vocal folds must also open and close between the mid-points of the preceding and following vowels. This took from 23 to 29 csec (see Klatt 1967). The results suggest that the consonants are placed as required by the intonation and stress patterns of the utterance; within this framework, articulatory movements follow each other as rapidly as phonetic circumstances permit.

Department of Phonetics
University of Leeds, Great Britain

REFERENCES

Björk, L.
1961 "Velopharyngeal Function in Connected Speech", *Acta Radiologica Supplementum* 202.
Klatt, D.H.
1967 "Articulatory Activity and Air Flow During the Production of Fricative Consonants", *M.I.T. Quarterly Progress Report* 84:257-260.
Scully, C.
1969 "Problems in the Interpretation of Pressure and Air Flow Data in Speech", *Phonetics Department Report* 2:53-92 (University of Leeds).
1971 "A Comparison of /s/ and /z/ for an English Speaker", *Language and Speech* (April-June).

DISCUSSION

LADEFOGED (Los Angeles)
Could you comment on whether your data agree with the data reported by Klatt and Stevens, who found that there was a greater rate of flow in voiced fricatives than in the adjacent vowel?

SCULLY
For the one speaker who acted as subject in the experiment, the air flow rate was lower in the voiced fricatives than in the adjacent vowel, which carried the sentence stress. In interpreting this data it does not seem to be essential to postulate an abduction of the vocal folds during the production of these voiced fricatives for this particular speaker of Southern English.

CATFORD (Ann Arbor, Mich.)
Re the example 'Say the fate afar', it was suggested that the air flow trace indicates incomplete closure for the /t/. Is this really so? Or is it an artifact, due to air flow generated IN FRONT OF the articulatory closure — which is quite common and is sometimes wrongly interpreted as indicating non-closure of the articulation.

SCULLY
Your comments raise very important problems of interpretation of aerodynamic data. I have discussed these problems elsewhere (Scully 1969), but dealing with them quantitatively is much more difficult. Experiments with a closed glottis are in progress to investigate air flow arising from cavity volume changes and articulator movements. One way to minimise this problem is to deliberately choose utterances which avoid large jaw movements and tongue forward-backward movements. But, in addition, cineradiography should be used in conjunction with air flow and pressure measurements to obtain estimates of the effects of particular volume changes upon the aerodynamic data.

Re the particular case of the /t/; model experiments in progress are indicating that turbulence noise intensity falls as the tongue constriction area is reduced below an optimum value, so that complete tongue closure may perhaps not be completely essential in all phonetic contexts.

CATFORD
Re Dr. Ladefoged's comment on the relative volume-velocity of air flow in voiced fricatives and vowels, my own experience is that this varies: sometimes the fricative has higher flow-rate, sometimes the vowel.

SCULLY
Re the comments of Prof. Catford and Prof. Ladefoged about glottal air flow in voiced fricatives and vowels. My experience agrees with that of Prof. Catford. For example, in my data, an unstressed /ə/ can have higher air flow rate than the /z/ which precedes it and I interpret this as a glottal opening movement which is taking place during the /ə/ in preparation for a following voiceless fricative.

The problems of interpretation already raised by Prof. Catford apply here also: in my data the stressed diphthong /eɪ/ sometimes shows zero or even negative oral air flow, probably because of large jaw movements. When deducing volume velocity of glottal air flow from data on volume velocity of air flow out of the mouth there are additional corrections to be made arising from air pressure changes inside the oral pharyngeal cavity and changes in the volume of that cavity, both active and passive. Also, of course, nasal air flow must be included in the calculation.

Categorical statements about glottal mechanisms will need to be based on more sophisticated aerodynamic and glottographic experimentation and more data than we have available at present.

STEVENS (Cambridge, Mass.)
If I read your figures correctly, the examples of air flow and turbulence noise for [t] and [s] were for utterances in which the consonants were in post-stressed position. Is it not possible that the apparent incomplete closure for [t], and the similarity in the rate of release for these two consonants is a consequence of this particular phonetic environment? One might not, perhaps, expect these properties to occur when the consonants are in prestressed position. I would anticipate under these circumstances a slower rate of increase of cross-sectional area for the fricative than for the stop.

SCULLY

I agree that aerodynamic and other data indicate complete tongue-alveolar ridge closure for /t/ in other contexts, for example before a stressed vowel. I do not know whether different stress contexts might automatically lead to different degrees of closure for the tongue or whether, alternatively, the tongue-blade trajectory might be deliberately adjusted to cope with different aerodynamic conditions such as, for example, rising or falling tracheal pressure.