RELEASE RATES FOR [t] IN VCV SEQUENCES ESTIMATED FROM AERODYNAMIC DATA

Sheila J.Mair, Department of Electronics and Computer Science,
The University of Southampton, Southampton SO17 1BJ, U.K.

Celia Scully, Speech Laboratory, Department of Psychology,
The University of Leeds, Leeds LS2 9JT, U.K.

ABSTRACT

According to Stevens [1], "quantitative data must be obtained on rates of release and closure of articulators". Here, we use aerodynamic data in an orifice equation to estimate the rate of increase in the cross-sectional constriction area for [t] in different vowel contexts for 10 English speakers. Analyses of the results indicate that in most cases, the rate of release of [t] is significantly faster when an open vowel follows than when a close vowel follows.

RECORDINGS

Some data were recorded in 1987/88 for 10 adult speakers of Received Pronunciation English as part of the Alvey Project MMI 009, Speech Pattern Algorithmic Representation, and are referred to henceforth as the SPAR database. The speakers are: HB, JH, SR, GB, EA (female) and JM, MA, DH, JW, MB (male).

The recording sessions were carried out in the Department of Linguistics and Phonetics at the University of Leeds. Four channels of data were recorded onto FM tape: sound pressure (microphone signal), laryngograph signal, volume flowrate of air, interpreted as oral airflow for non-nasal sequences (measured with a Rothenberg mask) and intraoral air pressure (measured with an orally-inserted polyethylene tube). The airflow and air pressure signals were low-pass filtered at $50Hz$ before being recorded onto a mingograph along with the other two (unfiltered) signals.

MEASUREMENTS

The speech material analysed formed part of Set C2F of the SPAR database. This consisted of repeated [pəCV] sequences where C and V stand for various consonants and vowels respectively. Sequences in which C = [t] and V = [iː, ɑː, ɔː, uː] were selected for analysis. Repetitions 2, 3, 4, 5 and 6 of each vowel context were analysed for each speaker. Measurements of airflow and air pressure were made at 10ms intervals following the plosive release, with the time of release defined from the rapid increase in flow from zero or near-zero. Using an orifice equation, the increasing minimum cross-sectional area of the vocal tract constriction is estimated. The equation is:

$$A_c = 0.00076 \times U_c/P_c^{0.5}$$

where A_c is the minimum cross-sectional area of the constriction (in cm^2), U_c is the volume flowrate of air through it (in cm^3/s), and P_c is the pressure drop across the constriction (in cmH_20); the orifice equation is discussed in more detail in Scully [2].

RESULTS

Graphs of constriction area against time are plotted. The graphs suggest that the increase in constriction area in the initial part of a [t] release is approximately linear, and that the release is faster in the open vowel contexts ([ɑː] and [ɔː]) than the close vowel contexts ([iː] and [uː]). As examples, graphs for [tiː] and [tɑː] are presented for Speaker HB in Figure 1.

Based on the area increase in the initial 50ms following the release, rates of release are calculated for each repetition. Release rates (with means and standard deviations) are presented for the different vowel contexts for each subject in Table 1.

A one-way analysis of variance indicates that there is a very highly significant effect of vowel context on the rate of release of [t] for all speakers except DH and EA ($p \leq 0.001$).

DISCUSSION

Of the vowels analysed, [ɔː] is likely to have most lip-rounding for Received Pronunciation speakers. Lip-rounding may begin during the consonant due to processes of coarticulation and so there may be a significant pressure drop across the rounded and protruded lips. In such a case, the pressure drop across the alveolar constriction may be less than the measured intraoral air pressure suggests. Therefore constriction area values calculated with the orifice equation for this vowel context may be under-estimating the actual values.

The release rates calculated here are generally consistent with the range of $5-20cm^2/s$ estimated by Fant [3] from acoustic analyses of formant transition patterns studied from spectrograms of plosives.

Massey [4] estimated a typical release rate of $100cm^2/s$ for labial and alveolar plosives (compared to $25cm^2/s$ for velar plosives). This value seems rather high compared to the results here, even for the open [ɑː] vowel context.

Measurements of X-ray data for [t]

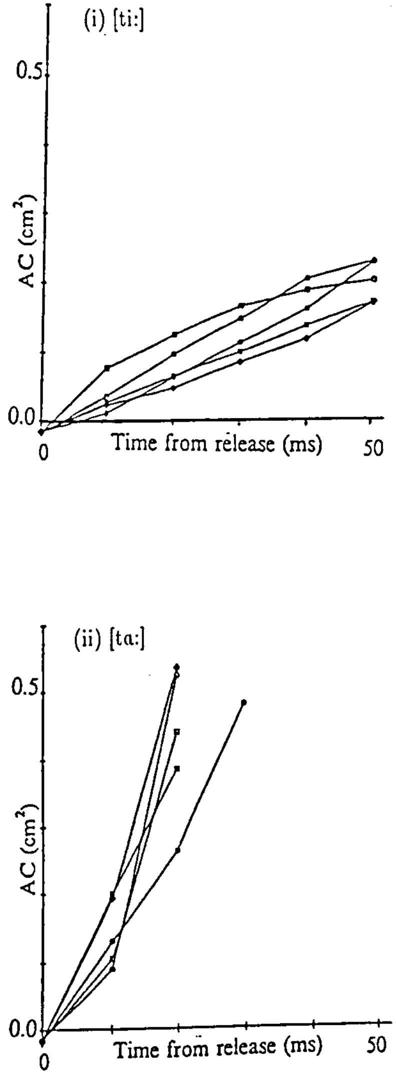

Figure 1: *Graphs of release of (i) [tiː] and (ii) [tɑː] for Speaker HB.*

Table1. Release rates (in cm^2/s) estimated from aerodynamic data for [t] preceding different vowels, for the 10 SPAR speakers.

	JM	MA	DH	JW	MB	HB	JH	SR	GB	EA
ti:2	2.5	5.4	5.1	2.6	7.1	4.9	3.6	3.8	4.4	2.7
ti:3	4.8	5.5	5.1	3.3	6.5	3.7	3.4	3.0	4.5	1.9
ti:4	4.9	6.7	5.7	3.1	7.8	3.7	3.6	2.9	4.3	2.6
ti:5	3.6	5.5	4.3	2.8	8.3	4.9	3.2	3.4	2.7	3.1
ti:6	4.1	6.4	3.3	1.1	7.1	4.3	3.5	3.2	5.4	3.0
Mean	4.0	5.9	4.7	2.6	7.4	4.3	3.5	3.3	4.3	2.7
St.Dvn.	0.98	0.60	0.93	0.87	0.70	0.60	0.17	0.36	0.98	0.47
tɔ:2	7.9	7.0	5.7	16.7	12.6	18.6	6.0	6.3	10.0	6.2
tɔ:3	9.6	8.6	5.7	13.8	10.9	19.3	6.3	7.3	7.1	5.4
tɔ:4	10.7	9.1	6.3	9.6	7.8	13.1	4.9	6.1	7.4	5.4
tɔ:5	10.4	8.7	5.3	7.3	11.3	15.4	12.8	7.9	6.9	4.3
tɔ:6	10.4	9.8	5.4	5.5	11.3	15.3	8.7	7.9	8.0	4.1
Mean	9.8	8.6	5.7	10.6	10.8	16.3	7.7	7.1	7.9	5.1
St.Dvn.	1.14	1.03	0.39	4.62	1.79	2.6	3.15	0.86	1.26	0.87
ta:2	9.0	12.2	5.4	2.1	12.1	16.1	24.6	7.3	14.4	7.3
ta:3	4.7	13.6	5.6	8.1	11.8	19.4	36.1	6.7	13.7	4.9
ta:4	10.8	14.6	2.8	9.0	17.0	26.9	18.1	6.8	10.7	3.7
ta:5	7.1	16.0	2.2	5.6	13.3	26.3	19.4	7.0	8.6	3.3
ta:6	2.8	11.4	3.9	7.5	13.0	22.1	21.6	8.2	8.1	5.7
Mean	3.9	13.6	3.4	6.5	13.4	22.2	24.0	7.2	11.1	5.0
St.Dvn.	3.21	1.84	2.05	2.74	2.08	4.58	7.22	0.60	2.87	1.61
tu:2	1.2	9.2	2.5	1.6	8.9	3.8	4.1	4.0	5.8	4.1
tu:3	1.1	8.3	4.3	1.5	8.1	4.7	5.7	4.3	5.6	3.7
tu:4	2.0	9.1	3.9	3.3	9.4	5.5	6.3	4.4	5.3	2.3
tu:5	0.6	9.0	4.0	3.7	12.0	5.7	7.5	3.5	5.2	3.1
tu:6	0.9	8.4	3.5	2.9	12.3	6.1	4.9	4.4	6.0	4.1
Mean	1.2	8.8	3.6	2.6	10.1	5.2	5.7	4.1	5.6	3.5
St.Dvn.	0.52	0.42	0.70	1.0	1.90	0.92	1.30	0.38	0.33	0.77

for a male speaker of North-American English have demonstrated that the velocity of tongue movement following consonant release is "dependent on the target configuration of the following vowel" [5]. Those articulatory data are consistent with our aerodynamically-derived constriction area estimates, which have suggested that the articulatory release of the English plosive [t] in VCV sequences is faster when an open vowel follows than when a close vowel follows.

CONCLUSIONS

In the orifice equation, the measured intraoral air pressure is actually the pressure drop across the constriction, the teeth and the lips, and so the results do not necessarily indicate an actual single constriction of the vocal tract. However, the constriction area estimates derived from the aerodynamic equation do indicate consistent effects for a [t] release in different vowel contexts (faster when an open vowel follows than when a close vowel follows) and these are likely to produce consistent effects in the corresponding acoustic signal.

The shape of the vocal tract constriction and its position along the vocal tract length will also have acoustic effects which are manifest throughout the transition to a following vowel [6]. Both these parameters are likely to vary for [t] in different vowel contexts. New methods for gathering articulatory data, such as enhanced electropalatography [7], could provide invaluable information about the three-dimensional shape of the vocal tract constriction.

Simultaneous recordings of articulatory, aerodynamic and acoustic data could help our understanding of the mapping between all these different aspects, and of the enormous complexities involved in speech.

ACKNOWLEDGEMENTS

Thanks are due to Eric Brearley in Leeds for his help with the data acquisition, and to the 10 SPAR speakers. This research was supported in part by a SERC studentship award.

REFERENCES

[1] Stevens, K.N. (1991) "The contribution of speech synthesis to phonetics: Dennis Klatt's legacy", plenary lecture in the *Proceedings of the 12th ICPhS*, Aix-en-Provence, Vol.1, pp.28-37.

[2] Scully, C. (1986) "Speech production simulated with a functional model of the larynx and the vocal tract", *Journal of Phonetics*, Vol.14, pp.407-413.

[3] Fant, G. (1970) *Acoustic Theory of Speech Production*, Mouton, The Hague, second edition, p.199.

[4] Massey, N.S. (1994) *Transients at stop-consonant releases*, unpublished Masters dissertation, Massachusetts Institute of Technology.

[5] Perkell, J.S. (1969) "Physiology of speech production: results and implications of a quantitative cineradiographic study", *Research Monograph No.53*, MIT Press, Cambridge, Massachusetts, p.16.

[6] Maeda, S. (1987) Articulatory-acoustic relationships in unvoiced stops - a simulation study", *Proceedings of the 11th ICPhS*, Tallin, Estonia, USSR, Vol.5, pp.11-14.

[7] Chiu, W.S.C. and Shadle, C.H. (1992) "Use of palate shape data in an enhanced electropalatographic system", *Proceedings of the Institute of Acoustics*, Vol.14, Part 6, pp.415-422.