AN ARTICULATORY STUDY OF THE WORD-INITIAL /stɹ/ CLUSTER IN BRITISH ENGLISH

Mitsuhiro Nakamura

Department of English, Aoyama Gakuin University, Tokyo, Japan
m_nakamura@cl.aoyama.ac.jp

ABSTRACT

This paper explores the articulatory characteristics of the word-initial /stɹ/ cluster in British English. Of specific interest is the phenomenon called /s/-retraction. Using EPG and EMA data collected from the MOCHA-TIMIT database, variable realizations of the /stɹ/ cluster were analyzed for four speakers. The tongue-palate contact pattern of the /s/ in the /stɹ/ cluster (e.g. *street*) was compared with that of the onset singleton /s/ and /ʃ/ (e.g. *seat*, *sheet*). The temporal extent of anticipatory labial coarticulation was inspected. The results reveal that (i) for one speaker, the contact anteriority index of the /s/ in the /stɹ/ cluster is significantly different from that of the onset singleton /s/, (ii) /t/-affrication (e.g. *street*>*s*[tʃ]*reet*) is found for all the speakers, (iii) the constriction peak of the lower lip gesture is systematically attained at the release of the /t/ in the cluster. These preliminary findings are discussed in terms of speaker-dependent articulatory strategies.

Keywords: /s/-retraction, EPG, EMA, MOCHA-TIMIT

1. INTRODUCTION

The phenomenon of /s/-retraction in the /stɹ/ cluster (e.g. *street* > [ʃ]*treet*) has been investigated in various accents of English: for instance, [3, 11, 13, 14] for American English, [1, 2] for British English, [15] for Australian English, [9, 17] for New Zealand English, [16] for a cross-dialectal analysis. The current study aims to further investigate the articulatory nature of the /s/-retraction phenomenon in British English.

Previous studies have identified that the phonetic basis of /s/-retraction is related to the lower spectral frequency: the alveolar /s/ in the /stɹ/ sequence has been judged to become closer to the postalveolar /ʃ/ due to its lower fricative noise frequency [2, 3, 11]. This measure successfully captures the systematic acoustic variation of /s/ in various contexts. In the current study, an attempt will be made to explore the 'articulatory' retraction of /s/ to /ʃ/ in the /stɹ/ cluster using electropalatographic data from the multi-channel articulatory database, MOCHA-TIMIT.

A previous study of /s/-retraction in British English [2] has proposed that /t/-affrication is the most important factor in modeling /s/-retraction in the /stɹ/ cluster and the rhotic /ɹ/ plays only an indirect role. According to [2], the retraction of /s/ occurs not only in the consonant sequence /stɹ/ (eg. *street*) but also in /stj/ (phonetically [stʃ] as in *student*). Since the consonant sequence /tj/ is affricated, the trigger for /s/-retraction is considered to be the affricate [tʃ]. This proposal effectively supports Lawrence's 'local assimilation' [9] as the mechanisms of /s/-retraction, rather than Shapiro's 'assimilation at a distance' [13].

The mechanism proposed by [2] is appealing but needs further examination. The first point concerns /t/ affrication mentioned above. It has been pointed out by [6, 10, 18] that the /tɹ, dɹ/ and /tj, dj/ clusters are realized as the affricates [tʃ, dʒ], but this tendency has not been substantiated fully by the empirical data. The second point is the coarticulatory effects of the rhotic /ɹ/. A study of the singleton /ɹ/ [5] has shown that the /ɹ/ involves more than one lingual constriction; the articulators are controlled in the order of the lips > the tongue tip/blade > the tongue root at the syllable onset position, and in the order of the tongue root-the lips > the tongue tip/blade at the coda position. Furthermore, long-range coarticulatory effects have been reported for the /l, ɹ/ in Southern British English [19]: adjacent and non-adjacent vowels preceding /ɹ/ have stronger lip rounding and tongue dorsum raising and retraction. It can be asked how the lip gesture is realized in the context of consonant clusters.

The current study aims to (i) provide initial empirical data for the /s/ articulation in the /stɹ/ cluster and (ii) examine the temporal extent of the lip gesture in the production of the /stɹ/ cluster. We will conclude with future directions of this study.

2. METHOD

2.1. Data collection

The speech materials used in this study were drawn from the multichannel articulatory database, MOCHA-TIMIT described by [20, 21]. This database comprises the articulatory and acoustic data of 460 phonetically balanced sentences read by native speakers of English. Data acquisition was made with Carstens Electromagnetic Articulograph (EMA, 500Hz sample rate), Reading Electropalatograph (EPG, 200Hz sample rate), Electrolaryngograph (Lx), and acoustic recordings. In the EMA data recording, the tongue tip (TT) coil was placed at about 5-10mm, the tongue blade (TB) coil was placed at about 25-40mm, and the tongue dorsum (TD) coil was placed at about 50-70mm from the tongue tip, as measured with the tongue extended. In this study, the EMA data and the EPG contact patterns were analyzed for the selected utterances spoken by four speakers of Standard Southern British English referred to as SE (female), JM (female), SA (male), and AP (male).

For collecting the speech materials, the MOCHA corpus was searched and two sets of data were prepared. Dataset 1 consisted of words with the word-initial /stɹ/ cluster: the target words were preceded by a schwa [ə] or non-coronal consonant. Dataset 2 consisted of words with the word-initial (onset) /s/ and words with the word-initial (onset) /ʃ/. These target words were preceded by a schwa [ə]. All the words collected are content words. The two datasets are summarized in Table 1.

Dataset 1	Dataset 2
street, straightforward, stray, stronghold, strongly	see, seeking, same, save, saw
	shipbuilding, shell, shaving, shocked, shore

Table 1: Target words in dataset 1 and 2.

The position in the sentence of these target words varied. The number of tokens in dataset 1 was 20 in total (5 words x 4 speakers) and that in dataset 2 was 40 in total (10 words x 4 speakers).

2.2. Analysis

The analysis of the EPG and EMA data was performed using a set of programs called EMA tools [21]. The Reading EPG artificial palate has 62 electrodes arranged in eight rows and eight columns. The surface of the palate was divided into two major zones for data analysis: an alveolar zone for rows 1-5 and a palatal zone for rows 6-8. These zones were subdivided into dentalveolar (row 1), alveolar (row 2), postalveolar (rows 3-4), prepalatal (row 5), medio-palatal (rows 6-7), and postpalatal (row 8).

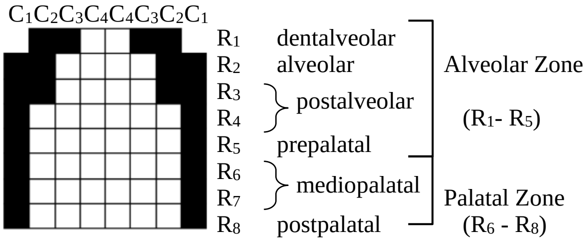

Figure 1: EPG palatogram and two zones.

The configurational characteristics of the EPG contact patterns in the alveolar zone were studied. For the /stɹ/ cluster (dataset 1), three EPG frames were identified. The MAX (linguopalatal narrowing) frame of /s/ was determined by the mid-point of the continuous noise energy, specified by a spectrogram. The two other frames were the first frame showing the complete closure for /t/ and the first frame showing the release of /t/. For the onset /s/ and /ʃ/ (dataset 2), the MAX frame mentioned above was specified.

The contact index method [7, 12] was used to examine the configurational characteristics of the EPG contact patterns. The contact indices were CAa (alveolar contact anteriority index), CCa (alveolar contact centrality index), and Qp (quotient of overall electrode activation at the palatal zone, rows 6-8). The indices CAa and CCa were calculated using the following formula [12].

(1) $CAa=[\log[[1(R_5/8)+9(R_4/8)+81(R_3/8)+739(R_2/8)+4921(R_1/6)+1]]/[\log(5741+1)]$

(2) $CCa=[\log[[1(C_1/8)+11(C_2/10)+121(C_3/10)+1331(C_4/10)]+1]]/[\log(1464+1)]$

The CAa index value increases with alveolar contact fronting. The CCa index was calculated for the four symmetrical columns (C_1 to C_4, C_4 to C_1 in Figure 1) and the CCa value increases as contact becomes more central in the alveolar zone. The Qp index was obtained by averaging all on-electrodes in the palatal zone by the total number of 24 electrodes. This value is related to the degree of tongue dorsum raising.

In order to explore the timing of the lip gesture on the /stɹ/ cluster, the EMA data of the lower lip movement (LLx, LLy) was examined. The time course of the lower lip gesture was identified by the four points (onset, target, release, and release offset) proposed by [8]. These time points were specified at 20% of the tangential velocity peak associated with movement towards, or away from, a target constriction. The total duration was the time interval between the onset and the release offset of a given movement and this was used to inspect the temporal extent of the lip gesture in the production of the /stɹ/ cluster.

The question to be explored in this study was which phase of the /stɹ/ cluster the constriction peak of the lower lip gesture was aligned to. For this purpose, the times of the tangential velocity minimum of the lower lip vertical movement (LLy) and the lingual gestures (the tongue tip, blade, and dorsum) were identified. Based on a qualitative observation, the interaction between the lip gesture and the lingual gesture was studied.

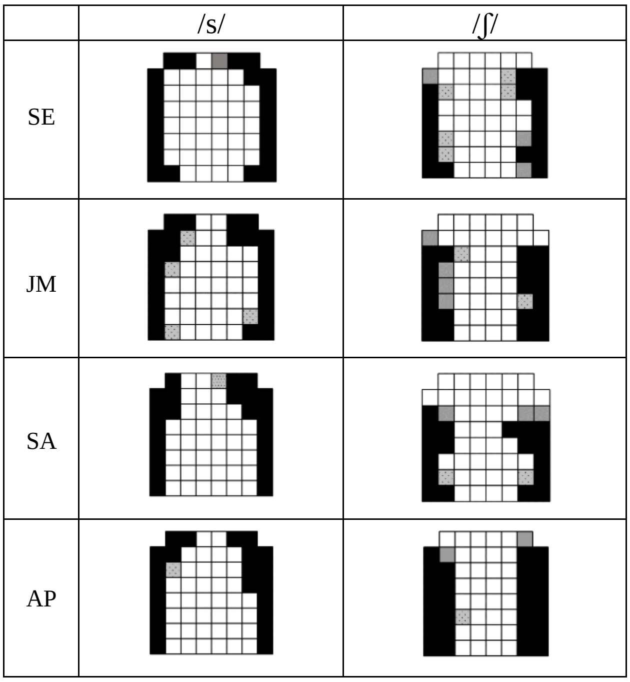

Figure 2: Averaged EPG contact patterns for the word-initial /s/ and /ʃ/. ('black' 100-80%; 'grey' 79-60%; 'dotted' 59-40%; 'white' 39≥)

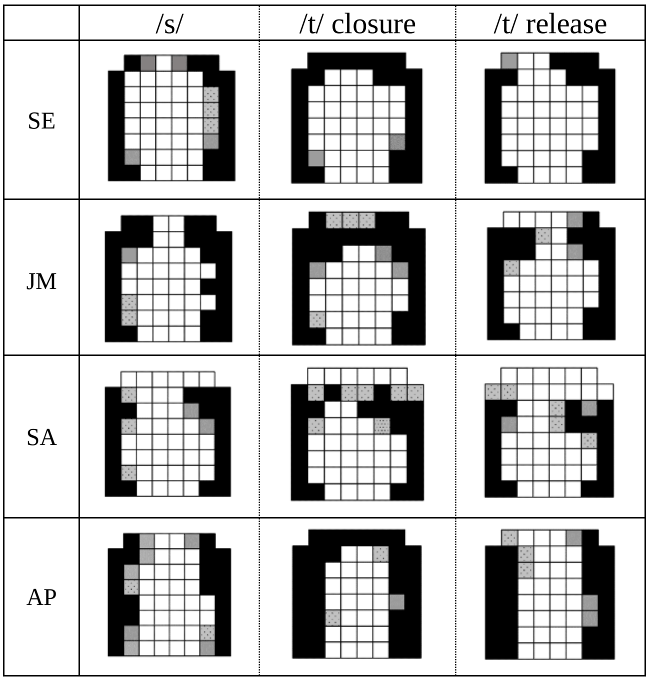

Figure 3: Averaged EPG contact patterns for the word-initial /stɹ/ cluster: /s/ at the MAX point, /t/ closure, and /t/ release. ('black' 100-80%; 'grey' 79-60%; 'dotted' 59-40%; 'white' 39≥)

3. RESULTS AND DISCUSSION

3.1. Linguopalatal contact patterns

Figure 2 above shows the EPG palatograms of the word-initial /s/ and /ʃ/ averaged over the productions of the target words. Figure 3 shows the EPG palatograms of the word-initial /stɹ/ cluster, which illustrate the MAX constriction of /s/, the closure of /t/, and the release of /t/, averaged over the productions of the target words.

The EPG patterns in Figure 2 serve as a basis for comparison. For the word-initial /s/, a narrow constriction is formed in the dentalveolar and alveolar regions (rows 1-2) with accompanying side contact. There are noticeable differences in the constriction formation among the speakers: apical (SE) and apicolaminal (JM, SA, AP) realization. For /ʃ/, the narrow constriction is primarily made in the alveolar and postalveolar regions (rows 2-4). This suggests that the tongue blade is involved in the constriction formation with the tongue tip directed downwards.

speaker	/s/	/ʃ/	/s/ in /stɹ/
SE	0.96 (0.01)	0.65 (0.04)	0.95 (0.02)
JM	0.96 (0.02)	0.54 (0.15)	0.95 (0.01)
SA	0.93 (0.01)	0.41 (0.15)	0.71 (0.01)
AP	0.95 (0.01)	0.77 (0.09)	0.93 (0.02)

Table 2: Mean values of the alveolar contact anteriority index (CAa) for singleton /s/ and /ʃ/, and /s/ in the /stɹ/ cluster. (s.d. in the bracket)

For /s, ʃ/ in the word-initial position and /s/ in the /stɹ/ cluster, the mean values of the contact indices are summarized: CAa in Table 2, CCa in Table 3 , and Qp in Table 4 (next page). To examine the differences between the singleton /s/ and the /s/ in the cluster, a t-test was carried out separately for the four speakers.

An examination of CAa showed a significant decrease only for speaker SA [$t(8)=25.22$, $p<0.01$]: SE [$t(8)=0.92$, $p=0.38$], JM [$t(8)=0.74$, $p=0.47$], and AP [$t(8)=1.93$, $p=0.08$]. Thus, the constriction location in the alveolar zone shifted considerably further back in SA's production.

In accordance with the results of CAa, a comparison of CCa revealed a significant decrease only for speaker SA [$t(8)=3.78$, $p<0.01$]: SE [$t(8)=0.06$, $p=0.95$], JM [$t(8)=0.85$, $p=0.41$], and AP [$t(8)=0.68$, $p=0.51$]. As seen in Figure 3, for SE, JM, AP, there is no substantial divergence between the constriction width for the /s/ in the cluster and that for the word-initial /s/. In contrast, for SA, the constriction width becomes wider at the (post)alveolar region (rows 2-3).

speaker	/s/	/ʃ/	/s/ in /stɹ/
SE	0.61 (0.11)	0.35 (0.08)	0.60 (0.12)
JM	0.57 (0.08)	0.33 (0.14)	0.54 (0.04)
SA	0.70 (0.01)	0.36 (0.09)	0.49 (0.11)
AP	0.48 (0.01)	0.34 (0.05)	0.47 (0.04)

Table 3: Mean values of the alveolar contact centrality index (CCa) for singleton /s/ and /ʃ/, and /s/ in the /stɹ/ cluster. (s.d. in the bracket)

speaker	/s/	/ʃ/	/s/ in /stɹ/
SE	33.33 (4.56)	41.67 (8.74)	46.67 (9.28)
JM	36.67 (7.17)	45.83 (3.73)	39.17 (6.24)
SA	27.50 (3.33)	38.33 (4.86)	34.17 (5.53)
AP	26.67 (2.04)	52.50 (5.65)	39.17 (11.67)

Table 4: Mean values of the quotient of overall electrode activation at the palatal zone, rows 6-8 (Qp) for singleton /s/ and /ʃ/, and /s/ in the /stɹ/ cluster. (s.d. in the bracket)

The results of Qp indicated a significant increase only for SE [$t(8)=-2.57$, $p=0.03$]: JM [$t(8)=-0.52$, $p=0.61$], SA [$t(8)=-2.06$, $p=0.07$], and AP [$t(8)=-2.11$, $p=0.06$]. As seen in Table 4, the increase in the palatal contact (rows 6-8) was found for all the speakers, and the significance of SE might be related to the production characteristic specific to the speaker: the apical articulation of /s/.

Based on the results of comparing the EPG indices, /s/-retraction was found for SA (characterized as a retractor) but not for SE, JM, and AP (characterized as non-retractors). Note that all the four speakers pronounced the /t/ in the /stɹ/ cluster as affricated.

It is worth noting that, while the non-retractors showed the EPG contact configuration of the /s/ in the cluster similar to the word-initial singleton /s/, they revealed idiosyncratic patterns at the /t/ closure and release phase. For JM, the contact at the frontmost row (row 1) tended to be lost: the complete closure for /t/ shifted backward (row 2); the constrictive approximation at the release phase of /t/ became similar to that of /ʃ/. For AP, the contact at both sides increased at the closure and release phase of /t/, and the contact configuration at the release phase was almost the same as the word-initial /ʃ/. For SE, the EPG contact at the closure and release phase of /t/ showed no substantial changes to /ʃ/: SE's production of the /stɹ/ cluster can be transcribed as [sts^{w}ɹ]. It is evident from our observation that when the 'articulatory' retraction of /s/ occurs in the /stɹ/ cluster, the affrication of /t/ also takes place, not vice versa.

3.2. Timing of the lower lip movement

Figure 4 (a, b) below shows the EMA recording of the tongue tip (TT), the tongue blade (TB), the tongue dorsum (TD), and the lower lip (LL) in the production of '*street*' in the MOCHA sentence, '*Spring Street is straight ahead*.' For speaker SA (a retractor) and SE (a non-retractor), the raising gesture of the lower lip extends over the whole /stɹ/ cluster (see the box in the LLy panel). The constriction peak of the lip gesture is attained at the release phase of the /t/. After the release of /t/, the lingual gestures (TTx, TBx, TDx) move smoothly to prepare the following /ɹ/. (For SE, no clear TD velocity minimum was observed.)

Such a coarticulatory accommodation of the lower lip was systematically observed across speakers and the /stɹ/ cluster in the target words. This could be considered as the effects of the rhotic /ɹ/. Another possibility is that, if the /stɹ/ sequence is realized phonetically as [stʃɹ], a postalveolar affricate might be accompanied by the lip rounding [4]. To understand these patterns fully, we need to analyze the rhotic gesture in various consonant sequences (e.g. /st/, /tɹ/, /spɹ, skɹ/), and to investigate the acoustic consequences of the lip gesture in the production of the cluster involving the /ɹ/.

4. CONCLUSION

The current study has explored the articulatory characteristics of the /stɹ/ cluster in British English. We have shown that (i) while /t/-affrication is observed for all the speakers, the 'articulatory' retraction of /s/ in the /stɹ/ cluster is found for one speaker only; (ii) the constriction peak of the lower lip gesture is systematically attained at the release phase of the /t/ in the sequence; and (iii) the non-retractors reveal idiosyncratic EPG contact patterns at the closure and release phase of /t/. These results provide evidence for speaker-dependent strategies and variability in the realization of /s/-retraction. Articulatory coordination in various clusters and its acoustic consequences will be substantiated by subsequent research.

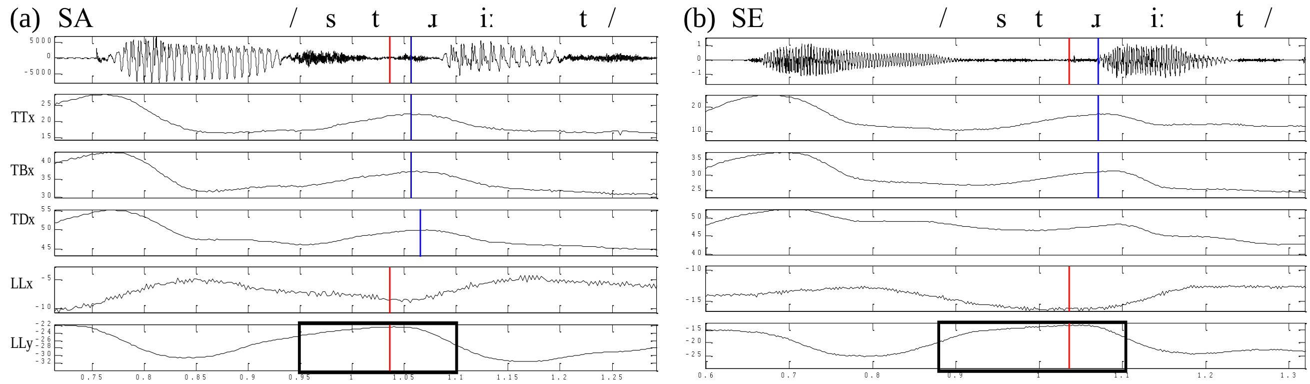

Figure 4: EMA trajectories of TTx, TBx, TDx, LLx, and LLy for (a) speaker SA and (b) speaker SE.
(red line=the constriction peak of LL; blue line=the constriction peak of TTx, TBx,TDx)

5. ACKNOWLEDGEMENTS

I would like to thank anonymous ICPhS reviewers for their critical and constructive comments. All remaining errors are of course mine. This research is supported in part by a grant from Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) 2023, Grant no. 21K00487.

6. REFERENCES

[1] Altendorf, U., MacDonald, R., Thielking, N. 2021. S Retraction in the South-East of England: A Snapshot of the Emergence of a Sound Change in Progress. *Anglistik: International Journal of English Studies* 32(1), 45-64.

[2] Bailey, G., Nichola, S., Turton, D., Baranowski, M. 2022. Affrication as the cause of /s/-retraction: Evidence from Manchester English. *Glossa: a Journal of general linguistics* 7(1), 1-30.

[3] Baker, A., Archangeli, D., Mielke, J. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23(3), 347-374.

[4] Brown, G. 1981. Consonant Rounding in British English: The Status of Phonetic Descriptions as Historical Data. In Asher, R.E., Henderson, E.J.A. (eds) *Towards a History of Phonetics*. Edinburgh: Edinburgh University Press, 67-76.

[5] Campbell, F., Gick, B., Wilson, I., Vatikiotis-Bateson, E. 2010. Spatial and Temporal Properties of Gestures in North American English /r/. *Language and Speech* 53(1), 49-69.

[6] Cruttenden, A. 2019. Gimson's *Pronunciation of English.* Eighth Edition. London: Routledge.

[7] Fontdevila, J., Pallares, M.D., Recasens, D. 1994. The contact index method of EPG data reduction. *J. of Phonetics* 22, 141-154.

[8] Gafos, A. 2002. A Grammar of Gestural Coordination. *NLLT* 20, 269-337.

[9] Lawrence, W.P. 2000. /str/→/ʃtr/: Assimilation at a Distance? *American Speech* 75, 82-87.

[10] Lindsey, G. 2019. *English after RP: Standard British Pronunciation Today*. Cham: Palgrave Macmillan.

[11] Mielke, J., Baker, A., Archangeli, D. 2011. Variability and homogeneity in American English /ɹ/ allophony and /s/ retraction. In Fougeron, C., Kuehnert B., D'Imperio, M., Vallee, N. (eds) *Laboratory Phonology 10*. Berlin: De Gruyter Mouton, 699-729.

[12] Recasens, D. Espinosa, A. 2007. An electro-palatographic and acoustic study of affricates and fricatives in two Catalan dialects. *J. of the IPA* 37, 143-172.

[13] Shapiro, M. 1995. A case of distant assimilation: /str/ →/ʃtr/. *American Speec*h 70, 101-107.

[14] Smith, B.J., Mielke, J., Magloughlin, L., Wilbanks, E. 2019. Sound change and coarticulatory variability involving English /ɹ/. *Glossa: a journal of general linguistics* 4(1), 1-51.

[15] Stevens, M., Harrington, J. 2016. The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English. *J. of Phonetics* 58, 158-134.

[16] Stuart-Smith, J., Sonderegger, M., Macdonald, R., Mielke, J., McAuliffe, M., Thomas, E. 2019. Large-scale acoustic analysis of dialectal and social factors in English /s/-retraction. In Calhoun, S., Escudero, P., Tabain, M., Warren, P. (eds) *Proc. 19th ICPhS*, Melbourne, 1273-1277.

[17] Warren, P. 2006. /s/-retraction, /t/-deletion and regional variation in New Zealand English /str/ and /stj/ clusters. In Warren, P., Watson, C.I. (eds) *Proc. 11th Australian International Conference on Speech Science and Technology, University of Auckland, New Zealand, December 6-8*, 466-471.

[18] Wells, J. 1982. *Accents of English.* 3 vols. Cambridge: Cambridge University Press.

[19] West, P. 1999. The Extent of Coarticulation of English Liquids: An Acoustic and Articulatory Study. In Ohala, J.J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A.C. (eds) *Proc.14th ICPhS*, San Francisco, 1901-1904.

[20] Wrench, A. 1999. MOCHA-TIMIT. https://www.cstr.ed.ac.uk/research/projects/artic/mocha.html

[21] Wrench, A. 2000. A multi-channel/multi-speaker articulatory database for speech recognition. *Phonus* 5, 1-13.