'Fax it up.' – 'Yes it does rather.' The relationship between the TRAP, STRUT, and START vowels in Aberystwyth English

Míša Hejná
Department of English, Aarhus University

Abstract

This paper focuses on the relationship between the TRAP, STRUT, and START phonemes in English spoken in Aberystwyth, mid Wales. It concludes that, in general, whilst the three phonemes do not show a complete merger, they do overlap to a great extent regarding a number of acoustic properties considered. It is predominantly F1 (for TRAP and STRUT) and vowel duration (for TRAP and START, and STRUT and START) that serve as correlates of the contrasts rather than F2 and the breathiness of the vowel. However, TRAP and STRUT do show a high degree of overlap in F1.

Unimodality tests reveal that TRAP and STRUT are associated with two independent modes regarding the distribution of F1 only in one of the ten speakers analysed. This does not apply to F2, vowel duration, and breathiness, where separate modes do not emerge in the statistical analyses for any speakers. On the other hand, a random forest analysis suggests that F1 is the most important correlate of the TRAP-STRUT contrast on the whole.

Regaring the START-TRAP and START-STRUT contrasts, unimodality tests show that START is associated with an independent mode concerning vowel duration for three of the ten speakers. A random forest analysis corroborates that vowel duration is generally the most important correlate of these two contrasts.

TRAP-STRUT merger

Crystal (2013) presents us with the following accent-related joke: "A judge arrives at his chambers having left an important document at home. 'Fax it up', his clerk suggests. 'Yes it does rather', replies the judge." Whilst this is presented as misinterpretation due to differences across accents, some researchers have noted a potential overlap between the TRAP and the STRUT vowels within a single accent. This has been suggested for RP by Wells (1982: 291-292), who comments on the increased perceptual similarity of the two phonemes in RP and who speculates that "[i]t may even be the case for some speakers that /æ/ and /ʌ/ are merged, variably at least." (1982: 292) Fabricius (2007: 311) has provided a quantitative production study that shows that TRAP and STRUT have indeed been changing regarding their absolute values in RP. However, Fabricius argues that these changes do not show configurational differences. Yet, some of the figures shown in Fabricius (2007) suggest a decrease in the difference between the phonemes on the F1 dimension.

Interestingly, RP is not the only accent for which a potential TRAP-STRUT merger has been proposed. An overlap between the two phonemes has also been reported for "some Southwestern Scots speakers" (Hall-Lew et al. 2017: 345). In addition, and crucially for the purposes of this paper, Hejná (2015: 271-3) has noted that there may indeed be a merger of the TRAP and the STRUT vowels (so that *fax* may sound like *fucks*) in English spoken in Aberystwyth, mid Wales (WE is used for Welsh English). These reports may imply a geographically relatively widespread potential merger within the UK.

Whilst Fabricius' work presents an essential instrumental study related to the subject-matter at hand, the incidental findings for Aberystwyth English, which point mainly to a merger in the F2 dimension, have not been examined in detail.

This paper therefore presents an analysis aimed to answer primarily the following question:

1. Is there a TRAP-STRUT merger in Aberystwyth English? More specifically, is there a merger in F1 and/or F2 of the TRAP-STRUT phonemes?

Although this study targets what may be a geographically limited phenomenon from a more global point of view (in comparison to, for example, High Rising Terminals – e.g. Britain 1992, Fletcher et al. 1999, Ritchart & Arvaniti 2014, Sando 2009, Shobbrook & House 2003, Shokeir 2008; GOOSE-fronting – e.g. Price 2008, Strycharczuk & Scobbie 2017, Ward 2003; and glottalisation – e.g. Cox & Palethorpe 2007: 342-343, Eddington & Channer 2010, Eddington & Taylor 2009, Foulkes & Docherty 1999, Mees & Collins 1999), any study of a potential merger enables us to explore the relationships between a number of acoustic dimensions relevant for vowel contrast beyond the traditional F1 and F2 dimensions. Importantly, in the studies of mergers, it has been pointed out that what may appear as full mergers when approached from F1 and F2 analyses may in fact constitute near mergers if the contrast is preserved via durational differences of the two vowels (see e.g. Labov & Baranowski 2006) or phonation (Di Paoli 1990). Links between phonatory settings and what is primarily thought of as oral phenomena have been shown and suggested (Brunner & Żygis 2011, Lotto et al. 1997); however, their insights remain to be extended to the field of sociolinguistics. This evidence motivates the following question:

2. Do vowel duration and phonation contribute to the phonetic implementation of the TRAP-STRUT contrast?

Answering the question will contribute to our understanding of the potential multiplicity of correlates of vowel contrasts, which have received less attention than consonants in this respect (e.g. Al-Tamimi & Khattab 2011, Toscano & McMurray 2010).

TRAP and START in WE

Whilst the START vowel has a backer quality in SSBE (Wells 1982), which makes it fairly distant from TRAP regarding F2, in WE it has been traditionally described as central (Wells 1982: 380-3), presumably due to the effects of Welsh. Additionally, the TRAP phoneme has been undergoing retraction in some British English accents (Wells 1982: 356, 380-383, 386). Considering that both TRAP and START can be central in WE, the following questions suggest themselves:

3. What are the phonetic correlates of the TRAP-START contrast?
4. What are the phonetic correlates of the STRUT-START contrast?

Methodology

This section introduces the social characteristics of the speakers, the structural properties of the material analysed, the technical aspects of the recording process, and finally the analysis of vowel formants, vowel duration, and breathiness.

Speakers

This study uses production data from 10 female participants, all of whom are native L1 Welsh speakers born and raised in Aberystwyth, mid Wales. The production data is based on their English production. At the time of the recording (2012-2013), the speakers' age ranged between 24 to 89 years. Age does not affect the results.

Data

The speakers read ˈCVC and ˈCVCV words, each type once in isolation (*hat* / *hut* / *heart*) and twice in a carrier sentence (*Say hat* / *hut* / *heart once*.). This was part of a larger project targeting fortis obstruents and the data analysed in this study represents only a subpart of the entire dataset. The segmental and the prosodic conditions are identical across the speakers. The second consonant is always a fortis obstruents (/p/, /t/, /k/ for all three vowels; /θ/, /s/, /ʃ/ for TRAP and STRUT; and /f/ for STRUT). In total, 1295 TRAP, 745 STRUT, and 471 START tokens were included in the present analyses. For more details, see Hejná (2015: 308-312).

Procedure

The speakers were reading the target words and sentences in a randomised order. They were wearing a head-mounted AKG C520 microphone and the recording device used was H4 Zoom Handy Recorder.

Statistical analyses were conducted in R (R Core Team 2018) and RStudio, using the lme4 (Bates et al. 2015), ranger (Wright et al. 2018, Wright & Ziegler 2015), and scatter3d (Ligger & Mächler 2003) packages.

Formant measurements

Formant measurements were extracted automatically in the midpoint of the modal interval of the vowel with the default 25ms analysis window in Praat (Boersma & Weenink 2014). The measurements included F1, F2, and F3, and were all checked manually for potential measurement errors. Different formant settings were used for the individual speakers, as appropriate. Tokens with whispered (i.e. voiceless) and exceedingly breathy vowels were excluded from the analyses (see Hejná 2015: 136 for more details).

Normalised values of the measurements were obtained via the NORM suite (Thomas & Kendall 2007-2014). Nearey formant intrinsic normalisation was used (Nearey 1977).

Durational measurements

The onset of the vowel and its offset were identified on the basis of the periodicity onset and offset through the inspection of the waveform. Vowel segmentation was conducted manually. Normalised vowel duration (quantified as a percentage of the overall word duration) provided the same results as raw vowel duration.

Breathiness measurements

The phonation characteristic of interest in the present study is that of breathiness, since the speakers produce glottalisation fairly infrequently in the data (excepting one of the speakers included in this study). CPP was therefore used to quantify the amount of breathiness in the vowels, as it has been identified as the best correlate of perceived breathiness (Hillenbrand, Cleveland & Erickson 1994: 776, Fraile & Godino-Llorente 2014). CPP measurements were extracted in VoiceSauce (Shue 2010; Shue, Keating, Vicenik & Yu 2011) in Matlab (2016).

Results

TRAP-STRUT formants

Pooling the data across all ten speakers, we find that the visual inspection of the distribution of F1 is bimodal (Fig. 1) and that of F2 is unimodal (Fig. 2). Hartigans' Dip Test (Hartigan & Hartigan 1985, Maechler 2015) nevertheless does not confirm the presence of a bimodal distribution in either case (F1: D = 0.005549, p-value = 0.9824; F2: D = 0.0061765, p-value = 0.9295).

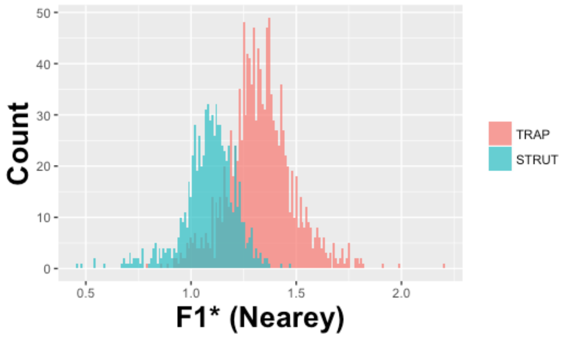

Figure 1. Distribution of F1 (normalized) for TRAP and STRUT.

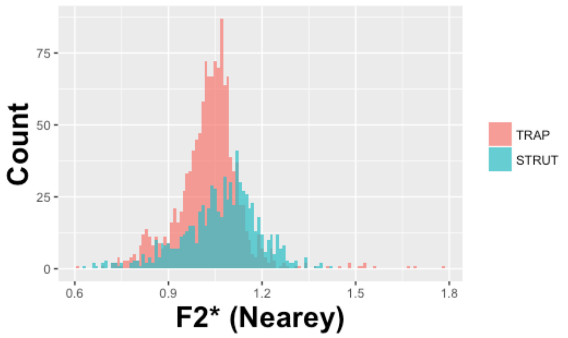

Figure 2. Distribution of F2 (normalized) for TRAP and STRUT.

Narrowing the analyses to the specific individuals, only one shows a bimodal distribution in F1 (ABE24; D = 0.04, $p < 0.01$), whilst the rest exhibit a unimodal pattern (D = 0.01-0.04, p = 0.11-0.996). Regarding F2, none of the speakers shows a bimodal distribution (D = 0.01-0.03, p = 0.45-0.995).

TRAP-STRUT duration

Pooling the data across all ten speakers again, we find that the distribution of vowel duration is unimodal (Fig. 3), and Hartigans' Dip Test confirms this unimodality (D = 0.01, p-value = 0.9922). A closer inspection of the individual speakers does not reveal any bimodality either (D = 0.02-0.03, p = 0.39-0.95).

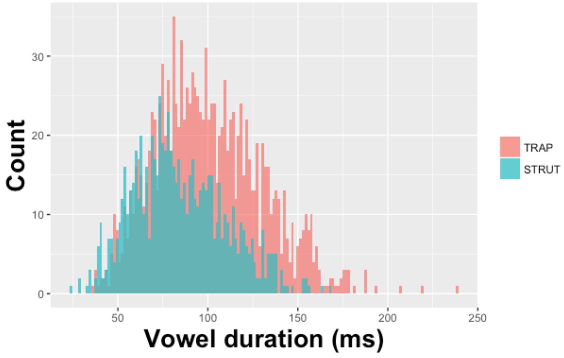

Figure 3. Distribution of vowel duration for TRAP and STRUT.

Place of START: formants

The distribution of F1 of the START vowel is found at the intersection of the STRUT and the TRAP phonemes (Fig. 4) and does not present a statistically confirmed distinct mode (D = 0.005, p = 0.98). Visually, START is associated with lower F2 than the other two vowel phonemes (Fig. 5), although this is not confirmed by the statistical analysis (D = 0.005, p = 0.97).

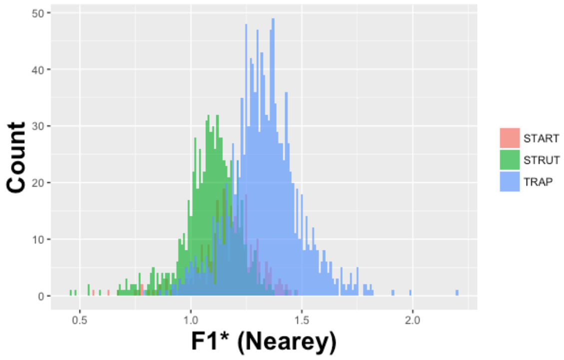

Figure 4. Distribution of F1 (normalized) for START with respect to TRAP and STRUT.

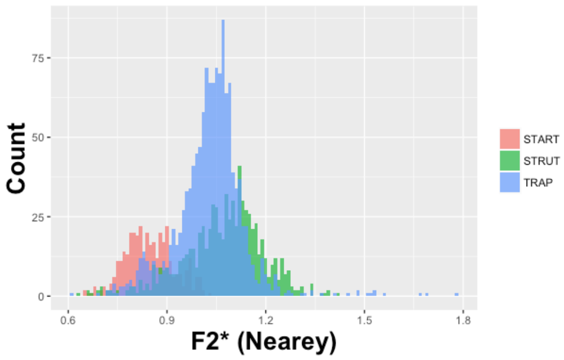

Figure 5. Distribution of F2 (normalized) for START with respect to TRAP and STRUT.

Individual variation is found only with respect to F1, for which ABE24 shows bimodality again (D = 0.03, p < 0.05). The remaining speakers suggest unimodal distributions (D = 0.01-0.03, p = 0.17-0.99). Considering F2, none of the speakers' distributions are bi- or multimodal (D = 0.02-0.03, p = 0.1-0.98).

Place of START: duration

Pooling all the speakers together, Fig. 6 shows that the START vowel is associated with fairly different durations than either STRUT or TRAP, although the Hartigans' Test rather surprisingly does not confirm this trend (D = 0.004, p = 0.99).

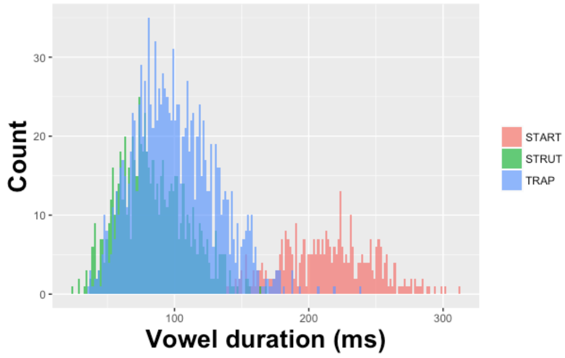

Figure 6. Distribution of vowel duration (ms) for START with respect to TRAP and STRUT.

Two individuals reveal a significant bimodal distribution (ABE31: D = 0.04, p < 0.05; ABE45: D = 0.04, p < 0.05) and one shows a bimodal trend approaching significance (ABE52: D = 0.03, p = 0.09). The remaining speakers do not exhibit bi- or multimodality (D = 0.02-0.03, p = 0.26-0.97).

Place of START: breathiness

Neither the visual inspection of the data, pooled across the speakers or by individual, nor statistical analyses reveal non-unimodality (D = 0.01-0.03, p = 0.71-0.998).

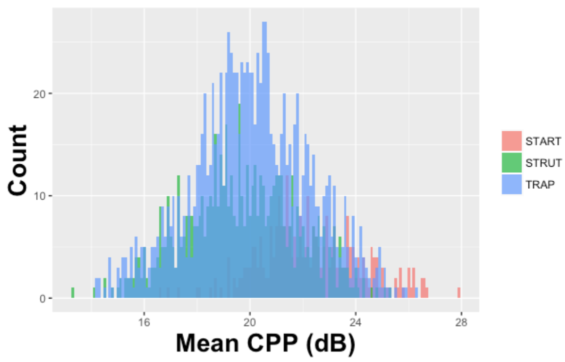

Figure 7. Distribution of mean CPP (dB) for START with respect to TRAP and STRUT.

Random forest analysis

Random forests analyses were used in order to estimate the importance of the four variables in distinguishing a. TRAP and STRUT, b. TRAP and START, and c. STRUT and START. The importance of the variables is summarised in Table 1: the higher the number, the more important the variable is in predicting the correct vowel phoneme identity (Gini importance index).

For the contrast between START and the other phonemes, vowel duration is without doubt the most important variable, following by F2 and F1 or the mean breathiness of the vowel (CPP). TRAP and STRUT, on the other hand, are again without doubt distinguished primarily by F1, following by F2, then vowel duration, and finally the mean breathiness of the vowel.

Table 1. Importance of the independent variables in predicting the three vowel contrasts; the higher the number, the more important the variable.

Pair	F1norm	F2norm	Vdur	CPP
TRAP-STRUT	424	132	85	67
STRUT-START	4	85	288	21
TRAP-START	23	87	328	16

Discussion

The analyses confirm that TRAP and STRUT are not distinguished by F2. However, the conclusion regarding F1 is more complex. On the one hand, the random forest analysis shows F1 as the most important correlate of the contrast, whose importance noticeably outperforms that of the other three variables considered. This suggests that the two phonemes have not merged regarding F1. On the other hand, the unimodality tests corroborate this result only for one of the ten speakers. Vowel duration and the mean breathiness of the vowel do not contribute to the phonetic implementation of the contrast in any obvious way, similarly to F2.

TRAP and START are distinguished primarily by vowel duration, as are STRUT and START. This is confirmed by unimodality tests for three of the ten speakers and more generally by the random forests analyses. Vowel duration emerges as the most noticeably important of the four variables considered in the implementation of the TRAP-START and STRUT-START contrasts. Although a number of studies have shown that vowel quality is essential in the perception of vowel contrasts of English, and indeed more important than vowel duration in the front part of the vocalic space at least (e.g. Bohn & Flege 1990), this presumably does not apply to the TRAP-START and STRUT-START distinctions, at least in Aberystwyth English. Interestingly, mixed effects modelling (not shown in this paper) provides rather different results, in which all four independent variables distinguish both of the contrasts considered here.

Ultimately, perceptual experiments need to help us reach conclusions as to whether there is a TRAP-STRUT merger in Aberystwyth English.

References

Al-Tamimi, J and G Khattab (2011). Multiple cues for the singleton-geminate contrast in Lebanese Arabic: acoustic investigation of stops and fricatives. *17th ICPhS, Hong Kong*: 212-215.

Bates, D, M Maechler, B Bolker and S Walker (2015). Fitting linear mixef-effects models using lme4. *Journal of Statistical Software*, 67, 1: 1-48.

Boersma P and D Weenink (2014). Praat: doing phonetics by computer. Version 5.3.78. <http://www.praat.org/> [accessed in 2014-2015].

Bohn, O-S and J Flege (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, 3: 303-328.

Britain, D (1992). Linguistic change in intonation: the use of high rising terminals in New Zealand English. *Language Variation and Change*, 4: 77-104.

Brunner, J and M Żygis (2011). Why do glottal stops and low vowels like each other? *17th ICPhS, Hong Kong*: 376-379.

Cox, F and S Palethorpe (2007). Australian English. *JIPA*, 37, 3: 341-350.

Crystal D (2013). *Language play*. Great Britain: Crystal ebooks.

Di Paolo M and A Faber (1990). Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language Variation and Change*, 2: 155-204.

Eddington, D and C Channer (2010). American English has goʔ a loʔ of glottal stops: social differentiation and linguistic motivation. *American Speech*, 85, 3: 338-351.

Eddington, D and M Taylor (2009). T-glottalization in American English. *American Speech*, 84, 3: 298-314.

Fabricius, A (2007). Variation and change in the TRAP and STRUT vowels of RP: a real time comparison of five acoustic data sets. *Journal of the International Phonetic Association*, 37, 3: 293-320.

Fletcher, J, E Grabe and P Warren (2004). Intonational variation in four dialects of English: the high rising tune. *Prosodic Typology. The Phonology of Intonation and Phrasing*. Jun, S-A (ed). Oxford: OUP.

Foulkes, P and G Docherty (1999). *Urban Voices: Accents Studies in the British Isles*. London: Arnold.

Fraile R and J I Godino-Llorente (2014). Cepstral peak prominence: a comprehensive analysis. *Biomedical Signal Processing and Control* 14: 42–54.

Hall-Lew, L, R Friskney and J Scobbie (2017). Accommodation or political identity: Scottish members of the UK Parliament. *Language Variation and Change*, 29: 341-363.

Hartigan, J A and P M Hartigan (1985). The dip test of unimodality. *The Annals of Statistics*, 13, 1: 70-84.

Hejná M (2015). *Pre-aspiration in Welsh English: a case study of Aberystwyth*. PhD thesis, University of Manchester.

Hillebrand J, R A Cleveland and R L Erickson (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37: 769-778.

Labov W and M Baranowski (2006). 50 msec. *Language Variation and Change*, 18: 223-240.

Ligges, U and M Mächler (2003). Scatterplot3d – an R package for visualizing multivariate data. *Journal of Statistical Software*, 8, 11: 1-20.

Lotto, A J, L L Holt and K R Kluender (1997). Effect of voice quality on perceived height of English vowels. *Phonetica*, 54: 76-93.

Maechler, M (2015). Package 'diptest'. Hartigan's dip test statistic for unimodality – corrected code. <http://cran.r-project.org/web/packages/diptest/index.html> [accessed in May 2015].

MATLAB 2016b, The MathWorks, Natick (2016).

Mees, I and B Collins (1999). Cardiff: a real-time study of glottalisation. *Urban Voices: Accent Studies in the British Isles*. Foulkes, P and G Docherty (eds). London: Arnold. 185-202.

Nearey T M (1977). *Phonetic feature systems for vowels*. PhD thesis, University of Alberta.

Price. J (2008). GOOSE on the move: a study of /u/-fronting in Australian news speech. *Interspeech 2008, Brisbane*: 346.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. <http://www.R-project.org/> [accessed in 2014-2015].

R Studio (2009-2013). Version 0.98.1049. <http://www.rstudio.com> [accessed in 2013-2015].

Ritchart, A and A Arvaniti (2014). The form and use of uptalk in Southern Californian English. Speech Prosody 2014, Dublin.

Sando, Y T (2009). Upspeak across Canadian English accents: acoustic and sociophonetic evidence. *Proceedings of the 2009 Annual Conference of the Canadian Linguistic Association*.

Shobbrook, K (2003). High Rising Tones in Southern British English. *15th ICPhS, Barcelona*: 1273-1276.

Shokeir, V (2008). Evidence for the stable use of uptalk in South Ontario English. *NWAV*, 36, 14, 2: 15-24.

Shue Y-L (2010). The voice source in speech production: data, analysis and models. PhD thesis, UCLA.

Shue Y-L, P Keating, C Vicenik and K Yu (2011). VoiceSauce: a program for voice analysis. *17th ICPhS, Hong Kong*: 1846-1849.

Strycharczuk, P and J Scobbie (2017). Fronting of Southern British English high-back vowels in articulation and acoustics. *JASA*, 142, 1: 322.

Thomas E R and T Kendall (2007-2014). NORM: The Vowel Normalization and Plotting Suite 1.1. <http://lvc.uoregon.edu/norm/> [accessed in January 2014].

Toscano, J C and B McMurray (2010). Cue integration with categories: eighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science*, 34, 3: 434-464.

Ward, M (2003). Portland dialect study: the fronting of /ow, u, uw/ in Portland, Oregon. MA thesis, Portland State University.

Wells J C (1982). Accents of English 2. The British Isles. Cambridge: CUP.

Wright, M N, S Wager and P Probst (2018). Package 'range'.

Wright, M N and A Ziegler (2015). ranger: a fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77, 1: 1-17.