This is the accepted version of the article, before final edits have been made. The complete and final version can be found using the details below

English World-Wide 35:3 (2014), 340–353. doi 10.1075/eww.35.3.04dru
issn 0172–8865 / e-issn 1569–9730

The role of orthography in (apparent) L2 dialect acquisition

Rob Drummond

Manchester Metropolitan University

Abstract

This paper investigates the role of orthographic representation in the production of the local variant of the STRUT vowel in the speech of Polish migrants living in Manchester. A previous study (Drummond 2013) showed that acquisition of the local variant depended on various social factors, yet this only took into account conversation data, leaving some word list data unexplored. Comparing the two data sets reveals a difference in degree of acquisition, with the wordlist data producing more of the local vowel. This is explained not as a case of dialect acquisition as such, but more as an effect of orthography triggering a connection to the L1 phonology, the particular nature of which mimics acquisition of the local variant.

Keywords

Orthography, L2 pronunciation, dialect acquisition, sociolinguistic variation, L2 phonology.

1. Introduction

This paper reports on one aspect of a much larger study (Drummond 2010) which looked at the extent to which Polish migrants who were living in Manchester, UK, acquired features of the local dialect. One of the features under investigation in the larger study was the STRUT vowel, of interest due to the fact that the local variant differs significantly from that used by the Polish speakers on arrival. A more focused discussion of the findings in relation to the acquisition of this particular feature can be found in Drummond (2013); however, only the data gathered through conversation were analysed, leaving aside the data produced from reading aloud word lists. When the two data sets are compared, it appears that the local variant is more likely to be produced when reading aloud than in conversation, an unexpected finding given the phonetic nature of the vowel in question (see below). This paper takes a closer look at the word list data, and explores the possibility that it is orthography rather than dialect acquisition that is affecting the production of the vowel.

The site of the study is Manchester, a post industrial city in the northwest of England, with a population of 483,000 (Office for national Statistics 2010); the participants are Polish migrants who cane to Manchester as adults after the expansion of the European Union in 2004.

1.1 *STRUT*[1]

The STRUT vowel is of interest here because in Manchester - and indeed in the north of England in general - there is no phonemic opposition between the STRUT vowel and the FOOT vowel, meaning both vowels are phonemically represented as /ʊ/. This lack of a so-called FOOT/STRUT split (Wells 1982:351-353) is in marked contrast to the pedagogical model of English the Polish speakers have been exposed to, a model which is based on the Standard Southern British English (SSBrEng) vowel system and which has a clear distinction between STRUT and FOOT, with STRUT phonemically

[1] Much of the following description of the STRUT vowel previously appeared in Drummond (2013)

represented as /ʌ/ (although a more accurate phonetic representation is the central [ɐ]) and FOOT remaining as /ʊ/. The result of this lack of split in the north of England is that words such as *put* and *putt* are homophonous for many local native speakers (NSs). This difference marks a salient distinction between northern and southern varieties of English. There is, however, a degree of variation amongst speakers in the realization of STRUT, even within the north. This is particularly true in the speech of those higher up the socioeconomic scale, where the STRUT vowel is often found to be intermediate between the two extremes ([ɐ] and [ʊ]), and to varying degrees. Wells (1982) discusses a few possible realizations for an intermediate sound, including a mid, central, unrounded [ə]. This is indeed the most common outcome of any STRUT variation in the speech of people in the Manchester area: a sound somewhere between (and including) [ʊ] and [ə] but almost never any more open than that.

1.2 *Acquisition of the local form*

In terms of acquisition of the local variant, the question here is the extent to which the representation of STRUT in the speech of the Polish speakers has moved from the pedagogical target [ɐ] towards the local variant of somewhere between [ʊ] and [ə].

Of course, in addition to the pedagogical vowel system and the local vowel system, the L1 vowel system is also relevant, particularly when we bear in mind the types of influence predicted by, for example, Flege's (1995) Speech Learning Model (SLM). The closest vowel to SSBrEng STRUT is Polish /a/, which is somewhat more open than the pedagogical target [ɐ]. According to the SLM, the proximity of these two vowels has the potential to cause difficulties, as the perception of the two would be very similar. Whether or not this is the case amongst the speakers involved in the present study is of interest, but does not interfere with the focus in terms of movement towards the local STRUT variant. This is because even if the original STRUT vowel in the speech of the Polish participants is slightly more open than [ɐ] due to influence from Polish /a/, movement towards [ə] and [ʊ] would still be as a result of local influence. That is to say, because the Polish influence is working in the opposite direction to the local influence, the two processes are very much separate and cannot be confused.

1.3 *Orthography*

The effect of orthography has not hitherto been studied in relation to L2 dialect acquisition specifically, although its role has been noted in L1 dialect studies to a degree, and in studies into L2 phonological acquisition more widely[2].

One of Chambers' (1992) eight principles of (L1) dialect acquisition is that 'Orthographically distinct variants are acquired faster than orthographically obscure ones' (p.697), explaining the observation in his data that the loss of T-Voicing in his Canadian to British subjects happened more rapidly than the change towards R-lessness. This is because the loss of T-voicing is reinforced by orthographical representations (city = [t]; giddy = [d]) yet R-lessness is contradicted by orthographical representations, where the letter 'r' is not pronounced.

Of the more established models of L2 phonological acquisition, only Best & Tyler's (2007) PAM-L2 (a revised version of Best's (1995) original Perception Assimilation Model) refers to the effect of orthography, suggesting that the categorisation of new sounds into the L2 phonology may be affected by the L1 phonological representation. However, there has been more work done in the area of orthographic depth - the extent to which a language's writing system deviates from one-to-one grapheme to phoneme correspondence (Van den Bosch et al. 1994). Different systems can be seen to exist on a continuum from transparent (straightforward one-to-one grapheme-phoneme correspondence) to opaque (less consistent grapheme-phoneme correspondences) (Erdener &

---

[2] For a thorough review of orthography and L2 phonological acquisition see Rafat (2011). A very brief overview will be supplied here.

Burnham 2005). Erdener & Burnham (2005) carried out a fascinating study whereby Australian (opaque L1) and Turkish (transparent L1) speakers were tested with repetition tasks involving Irish (opaque) and Spanish (transparent) language items with and without orthographical stimuli. They found that the Turkish speakers were more affected by orthographic information than the Australian English speakers, illustrated by the fact that they made fewer mistakes in Spanish, and more mistakes in Irish, than the Australians, who were consistent across both languages. They argue that the Turkish speakers' reliance on orthography, to both positive and negative effect, is due to the transparency of their L1. Australians, with their opaque L1, have a different relationship with orthography that does not have such an effect in an L2 context. Polish is an example of an orthographic system that is nearer the transparent end of the continuum.

2. Methodology[3]

The participants for the study consisted of Polish adults who had grown up in Poland but were now living in Manchester. They all fulfilled the following criteria:

1. They grew up in Poland and came to England as adults
2. They were aged between 18 and 40
3. They had at least a basic proficiency in English before coming to England

In addition, ideal participants had lived nowhere else in the United Kingdom apart from the Manchester area. This was true for all but three participants. The final sample consisted of 40 individuals from a variety of backgrounds who were aged between 19 and 37. They were evenly split between male and female, and had been in the UK for between two months and six years. They were recruited through a variety of methods (contacts in the Polish community, social media, flyers in Polish shops etc) but the most effective method was word of mouth recommendation as the study progressed In addition to the Polish participants, four local native speakers were included to serve as

---

[3] Sections of the description of methodology previously appeared in Drummond (2013).

some kind of reference for the vowels. This sample consisted of one speaker aged 20-21 and one aged 35-40 of each gender, with all four having been brought up in the Manchester area.

Meetings were arranged with individuals throughout 2009 and speech data were gathered by way of an informal conversation, a picture task, and a wordlist with a view to exploring the effect of task formality on the speech of the participants. The focus here is on the wordlist data, with conversation data used for comparison purposes. Each meeting was recorded using a Zoom H2 Handy Recorder placed unobtrusively on a surface near the participant. Recordings were made as .wav files using a 44.1 kHz sampling rate with 16-bit precision, saved onto an SD memory card then transferred onto a PC.

All participants except one (speaker 3) took part in the wordlist, so the data presented here are from 39 participants. There is some variation in the number of tokens for each speaker, and this is the result of several factors. Firstly, two versions of the wordlist were used. Five speakers were given an earlier version of the word list which contained only six STRUT words, compared to a later version which contained eleven. The first version was felt not to offer a sufficient balance of words across the different variables under investigation in the larger study, so was changed. Secondly, some speakers, despite being urged to slow down, read the word list so quickly that the words could not be viewed as being good examples of isolated words. As one of the primary purposes of the word list was to get words in isolation, it was decided to discard these examples. Thirdly, there were a few occasions where a speaker did not know some of the words. It would have defeated the object of the task for the interviewer to help them, so these words were omitted. In practice, these final two reasons had relatively little effect – the five speakers who read from the earlier version each produced 5 or 6 usable STRUT tokens, and the remaining 34 speakers who read from the later version each produced 10 or 11 usable STRUT tokens.

2.1 *Coding STRUT*

Every STRUT token was coded into one of ten categories. Vowels which were auditorily perceived to be within the target NS range were coded in the first five categories 0-4 ([ɐ], [ɐ̹], [ə], [ʊ̹], [ʊ]), and vowels which were perceived to be outside the NS target range were coded in categories 5-9 ([ɔ], [ɒ], [u], [a], [ɑ]). The non-target vowels could be viewed as pronunciation errors, as they are not in line with either the pedagogical model or the local variety. Due to the fact that these tokens do not play a central part in the study, the categorizations are not necessarily intended to be phonetically precise. Instead, the symbols used represent the nearest cardinal vowel to what was heard, without the use of diacritics. Discussion of the 'five categories' below refers to the five NS target variants.

The five target variants actually lie on a continuum between the two extremes, so the categories are in some ways arbitrary, but categorizing them in this way helps to make sense of a continuous variable of this kind (Milroy & Gordon 2003). The decision to use five auditory categories was based on a process of trial and error during which it was found that the researcher could reliably distinguish more than three variants, yet not as many as six or seven. These auditory categorisations were backed up by a following acoustic analysis using Praat (Boersma & Weenink 2010) in which tokens were plotted, labelled with their auditory categorisation, and visually checked for consistency. Where individual tokens showed inconsistency between the auditory categorisation and acoustic measurement (even after re-checking) the auditory categorisation took priority, given the centrality in the study of the idea of perception. A more complete discussion of this process can be found in Drummond (2010; 2013). Categories 0 and 1 ([ɐ], [ɐ̹]) are viewed as representing unaffected, pedagogical standard Southern British English (SBrEng) vowels, and categories 2-4 ([ə], [ʊ̹], [ʊ]) are viewed as vowels having been influenced by Northern British English (NBrEng). In some ways it is perhaps useful to consider the variation as bipartite between these two categories, however, this arguably has the effect of over-simplifying the situation, and ignores the variation within each category.

3. Results and discussion

Initial comparisons between the dataset from the wordlist element and the dataset from the conversation element with regard to STRUT are striking. Figure 1 shows the distribution of target (NS target range, see above) tokens following auditory analysis for all speakers in both the conversation data (top chart) and wordlist data (bottom chart). Both charts are ordered by the mean auditory value across all five categories, although this results in a different speaker sequence for each chart. In both cases the four bars on the right represent the four native speakers. The first thing to note is that while only one Polish speaker exhibited no '0' ([ɐ]) tokens in the conversation element, this increased to four speakers in the word list. In fact, this reflects the overall tendency of the comparison – that the word list task produced more NBrEng influenced tokens than the conversation task. This can be seen to an extent in Figure 1; notice how the right hand side of the chart is generally darker in the bottom example, showing more local variants. However, Figure 2 provides a very clear illustration of the difference.

FIGURE 1 HERE

Figure 2 shows the difference between the mean STRUT auditory values for each speaker in the conversation element and in the word list. A positive difference shows an increase in the mean auditory value (thus, an increased use of NBrEng influenced variants in the word list) and a negative difference shows a decrease. Clearly, the majority of speakers show a move towards the local variant in the wordlist, with 20 of the 39 Polish speakers exhibiting a higher STRUT auditory value. Note, however, that all four NSs (darker grey) also show a move in the same direction

FIGURE 2 HERE

It is not immediately clear why this should be the case. It is, arguably, unlikely to be an example of a movement towards reduced articulatory effort, thus encouraging the STRUT vowel to be realised more centrally (and therefore more towards the NBrEng variant when starting from SBrEng [ɐ]), as this would be more likely to occur in the more rapid speech of the conversation task (although see below for a discussion of this possibility). One possible influencing factor is that of orthography, which will be discussed in due course.

In addition to the overall difference described above, it is also possible to look at the pronunciation of individual words from the list, with some words appearing more likely to encourage the use of a NBrEng influenced variant. Figure 3 shows the mean STRUT auditory value for each word[4] in the wordlist, along with the number of tokens for each (in brackets).

FIGURE 3 HERE

While it is clear that words such as *blood* and *hut* are more likely to show a variant closer to NBrEng STRUT than words such as *understood* and *mother*, the reasons behind the difference are not immediately obvious. A possibly lexical frequency effect was investigated by using data from the British National Corpus frequency lists supplied in Leech, Rayson, and Wilson (2001) and calculating the frequency of each word; however, there appears to be no correlation between the mean auditory STRUT value for the individual words and lexical frequency. Table 1 shows each word listed in order of auditory STRUT value (highest to lowest) with each word's BNC frequency. The lack of apparent correlation is confirmed when the BNC value is normalized using the log10 transformation and a Pearson correlation coefficient is calculated *(r=-0.218 p=0.520).*

TABLE 1 HERE

[4] *Starbucks* and *understood* are slightly different from the other words in the list, as the STRUT vowel does not fall on the primary stressed syllable of the word. However, both words tend to retain the full vowel in these syllables, which was certainly the case in all examples from the Polish speakers. For this reason they were included in the analysis.

While this pattern remains unclear, the same is not true when we look at non-target (outside the range of NS variation) realizations for each word. Recall that non-target STRUT variants were identified in terms of the closest cardinal vowel, which led to five possibilities: [u] [ɔ] [ɒ] [ɑ] [a]. Figure 4 shows the proportion of non-target STRUT realizations for each word, along with the number of tokens (target and non-target) of each. Notice that *blood* is again at the top of the list, meaning it has the highest mean auditory STRUT value (therefore closest to the local variant) as well as the highest rate of non-target realisations.

FIGURE 4 HERE

Although there is once again no correlation between the non-target proportion and BNC frequency (*r=0.301 p=0.368)*, there is perhaps evidence of the influence of orthography. The four words which show the highest rates of non-target vowels, *blood son money mother*, and indeed which have rates of non-target forms of at least three times those of all the other words, all have only the letter 'o' in their spellings for the STRUT vowel. It could be argued that it is this 'o' spelling of the vowel which is leading to the non-target realizations. This hypothesis is strongly supported when the details of the non-target forms are explored. Figure 5 shows each of the 9 words in which non-target variants were used, ordered from the most to the least tokens. Notice how in each of the four words with a 'o' or 'oo' spelling of STRUT, there is a high rate of [ɔ] and [ɒ] tokens. In fact, all the non-target tokens for *son money mother* are one of these two possibilities. If we then look at the Polish vowel system (Figure 6) we see that the Polish letter 'o' represents a vowel slightly below [ɔ]. This is unlikely to be coincidental, rather, it is an illustration of how an L1 grapheme/sound correspondence can be mapped onto the L2 system. This explanation is made more likely when we consider the nature of these particular L1 and L2 orthographic systems, in particular, their orthographic depth. As mentioned earlier, Polish is nearer the transparent end of the continuum, while English is nearer the opaque end. It is feasible, therefore, that the transparency of the L1 system is influencing the production of the L2 when faced with the (opaque) written form.

FIGURE 5 HERE

FIGURE 6 HERE

Orthography might also explain the finding that *blood* has the highest auditory STRUT value. Its high rate of non-target realisations (48%), and the fact that these non-target forms cover the widest range of possibilities (Figure 5) suggest that it is an unfamiliar word for several speakers. Unfamiliar, that is, in terms of reading; the word itself is not particularly unusual, but the spelling is. More importantly, it shares its spelling with other words which are more common in the FOOT lexical set, such as *look* and *good*. It might be the case that some of the tokens from *blood* which were auditorily categorized as '3' and '4' ([ʊ̜] and [ʊ]) were, in a sense, non-target realizations which just happened to match a target variant, thus falsely inflating the word's auditory STRUT value. In other words, *blood* was read as having a FOOT vowel due to its unfamiliarity, thus sharing a vowel with words such as *good* and *cook,* also in the word list.

There is also the possibility that orthography has an influence in the finding that the wordlist generally produced more NBrEng influenced (target) variants than the conversation task. The results above suggest it is likely that reading the words results in the participants being more aware of their spelling. Apart from the four words just described, the STRUT vowel in all the words is spelt with 'u' (or 'ou'). In Polish, the letter 'u' represents the vowel [u], which of course is very close to NBrEng [ʊ]. It might be the case that seeing the letter 'u' in the spelling of a word, rather than encouraging a pedagogical standard [ɐ] as one might expect, instead triggers a connection with the Polish vowel, thus colouring what is produced. The mechanisms of L2 proficiency restrict this (i.e. knowledge of English sound/spelling relationships gained through experience), preventing the vowel being realised as [u], yet the influence remains. The same influence is not at work in spontaneous speech, as the visual cue is not present. This explanation is made more likely when we recall that Polish has a more transparent orthography than English, and reflect on the results of Erdener & Burnham (2005) described earlier.

Orthography does not, however, explain the fact that the four NSs also exhibited a tendency to produce isolated words with 'stronger' NBrEng variants. Instead, this tendency might be explained by the point made earlier about increased articulatory effort in the wordlist data, resulting in vowels becoming more central in the conversation data, and more 'northern' in careful speech. Unlike most of the Polish speakers whose STRUT vowel would move closer to NBrEng as a result of reduced articulatory effort in the conversation speech (from [ɐ] to [ə]), the NS's STRUT vowel, by starting at something close to [ʊ], would actually move towards [ə] as a result of reduced articulatory effort. A corresponding reduction did not appear to occur in the spontaneous speech of the Polish speakers, as is evidenced by the direction of the difference between the two styles, most probably as a result of the more measured and monitored nature of L2 speech production.

The idea that the NNS wordlist vowels are more 'northern' for the same reason as the NS vowels (i.e. more careful articulation) is possible, but unlikely given the generally low levels of NBrEng variants overall. This can be seen in Figure 7, which shows the mean auditory STRUT value for each speaker in both tasks, ordered by degree of difference between the two speaking styles, but with the NSs separated on the right. For a similar process to be taking place between the NNSs and the NSs, we would expect the conversation STRUT value to be 2 and above, i.e. into the NBrEng region of [ə] - [ʊ], with the word list STRUT value higher. This would then reflect a pattern of reduced articulatory effort in the conversation (a movement towards [ə] from something more raised) and increased articulatory effort in the wordlist (a movement towards [ʊ̜] - [ʊ] from [ə]). If the conversation STRUT value is not 2 and above, then arguably it makes little sense to consider the process as being the same, as the articulatory effects would be working in reverse. Figure 7 shows that only 3 speakers have a conversation STRUT value of 2 or over, and only 2 of these (speakers 26 and 29) are likely candidates to be viewed in the same way as NSs. Clearly, the range of detail in the differences suggests more than one influence, and there may well be individual speaker differences that have not been addressed here, but overall, there is an argument for orthography playing a role in the patterning we see here.

FIGURE 7 HERE

4. Conclusion

Despite the data coming from a study into dialect acquisition in a second language, the results presented here do not, after all, add to that particular area of study. They do, however, illustrate the strength of the influence of orthography on pronunciation, particularly between languages, and particularly when those languages have different degrees of orthographic depth. Although apparently showing a process of dialect acquisition by showing a more advanced process of phonological acquisition than was evident in conversational speech, the differences in the word list data can perhaps be more convincingly explained by the effect of orthography, which triggers an L1 relationship that does not exist in conversational speech, thus mimicking rather than accurately representing the acquisition of the local variant. This finding perhaps helps highlight the continued importance of considering the role of orthography when researching issues of pronunciation.

References

Best, Catherine. 1995. A direct realist view of cross-language speech perception. *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. W. Strange (ed.). Timonium, York Press**:** 171-204.

Best, Catherine, & Michael Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), Second language speech learning: The role of language experience in speech perception and production (pp. 13-34). Amsterdam: John Benjamins.

Boersma, Paul and David Weenink (2010*). Praat: doing phonetics by computer*. [Computer program]. Version 5.1.17, retrieved 2009 from http://www.praat.org

Chambers, Jack. 1992. Dialect Acquisition. *Language* **68**(4): 673-705.

Drummond. 2010. *Sociolinguistic variation in a second language: the influence of local accent on the pronunciation of non-native English speakers living in Manchester.* Unpublished PhD thesis, University of Manchester.

Drummond. 2013. The Manchester Polish STRUT: Dialect Acquisition in a Second Language. *Journal of English Linguistics.* 41, 1: 65-93.

Erdener, V. Dogu and Denis K. Burnham. 2005. The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning* **55**(2): 191-228.

Flege, James Emil. 1995. Second language speech learning: theory, findings and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. W. Strange (ed.). Timonium, York Press**:** 233-272.

Jassem, Wiktor (2003). "Illustrations of the IPA: Polish." *Journal of the International Phonetic Association* 33(1): 103-107.

Leech, Geoffrey, Paul Rayson, et al. 2001. *Word frequencies in written and spoken English: based on the British National Corpus*. Harlow, Longman.

Milroy, Leslie and Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Oxford, Blackwell Publishers.

Rafat, Yasaman. 2011 *Orthography-induced Transfer in the Production of Novice Adult English-speaking Learners of Spanish*. Unpublished PhD Thesis, University of Toronto.

Statistics, Office for National. 2010. *Mid year population estimates 2009*.

Van den Bosch, Antal, Alain Contenty, et al. 1994. Measuring the complexity of writing systems. *Journal of Quantitative Linguistics* **1**(3): 178-188.

Wells, John. 1982. *Accents of English Vol 2*. Cambridge, Cambridge University Press.

Figure 1: Bar chart showing distribution of target STRUT tokens for all speakers (auditory analysis). Conversation element above, wordlist element below.

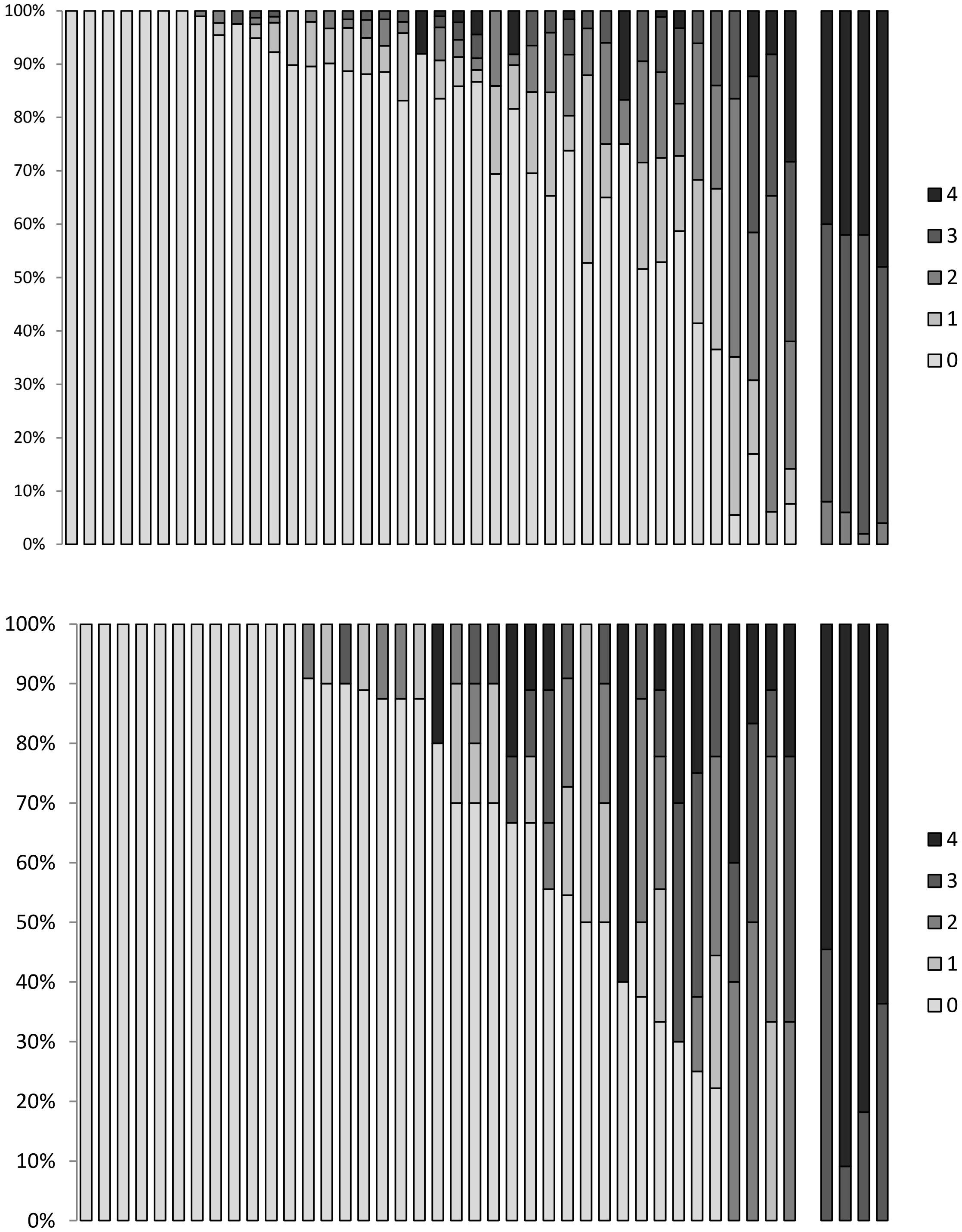

Figure 2: Chart showing the difference in mean STRUT auditory values between the conversation and wordlist elements.

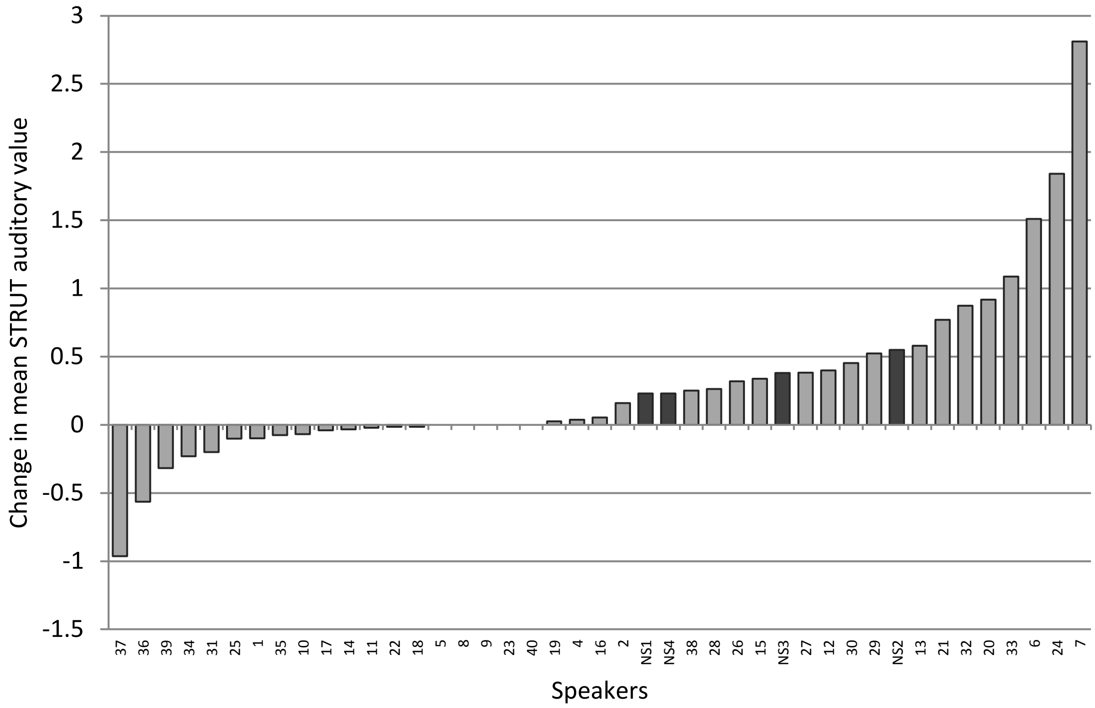

Figure 3: Chart showing the mean STRUT auditory values for each word.

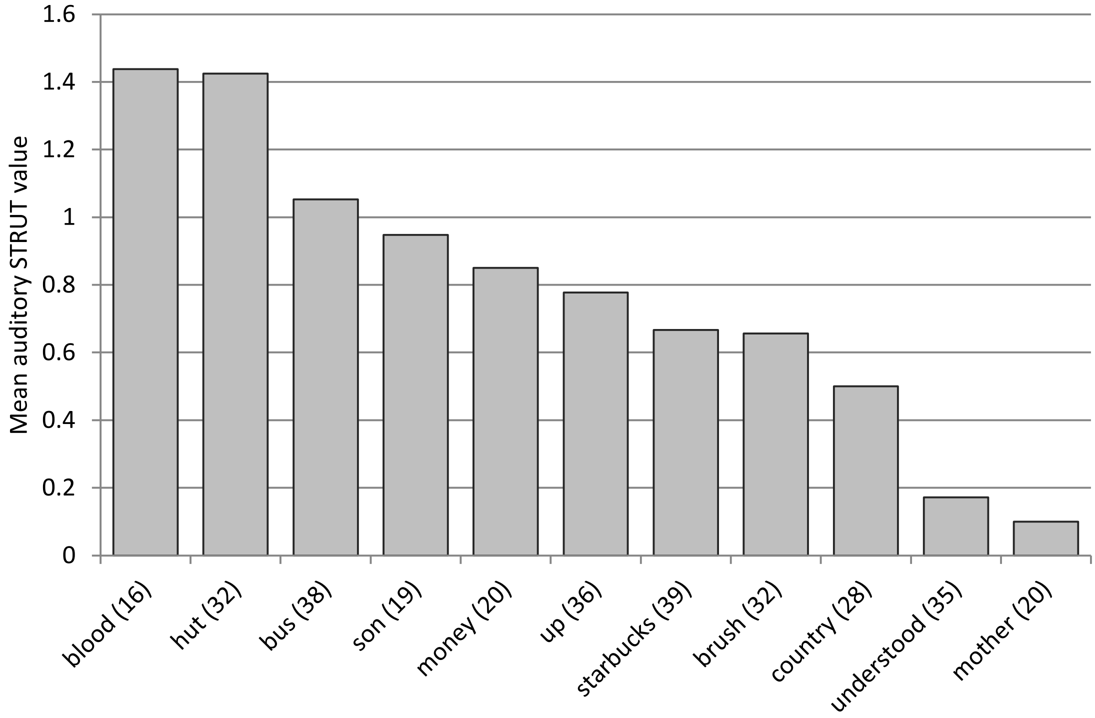

Table 1: Wordlist items with corresponding BNC frequency.

Word (in order of STRUT value)	BNC frequency per million words
blood	51
hut	<10
bus	94
son	72
money	637
up	3042
starbucks	<10
brush	13
country	204
understood	23
mother	184

Figure 4: Proportion of non-target STRUT realizations for wordlist items.

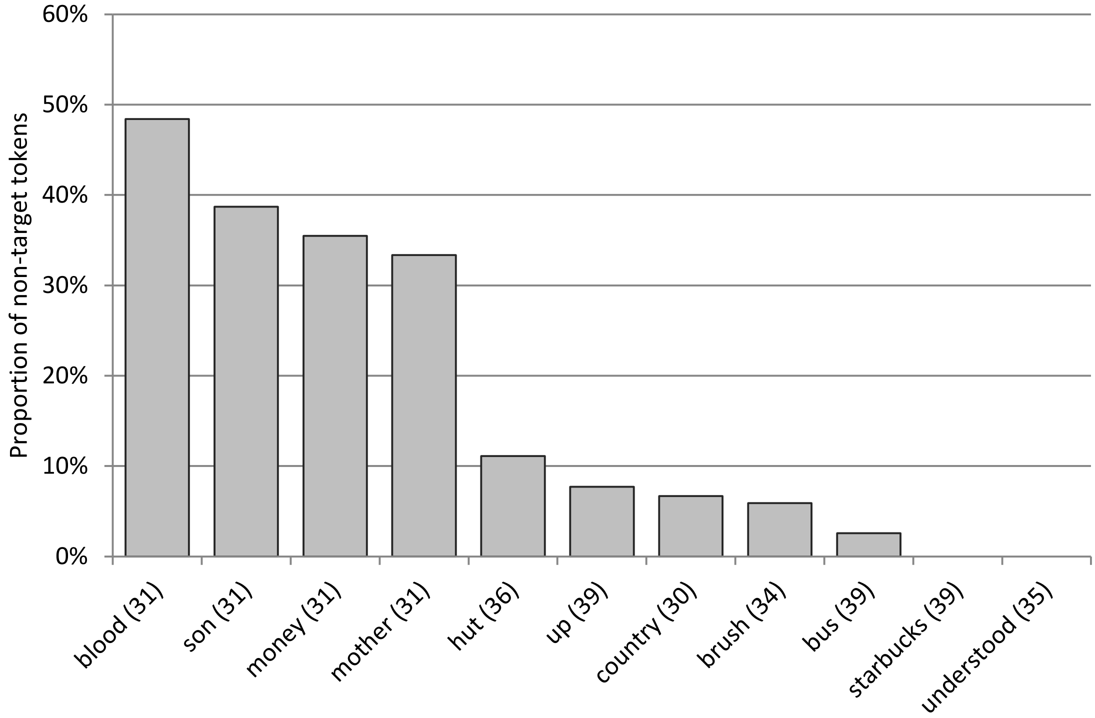

Figure 5: A breakdown of the non-target realizations of STRUT, ordered by proportion of non-target tokens.

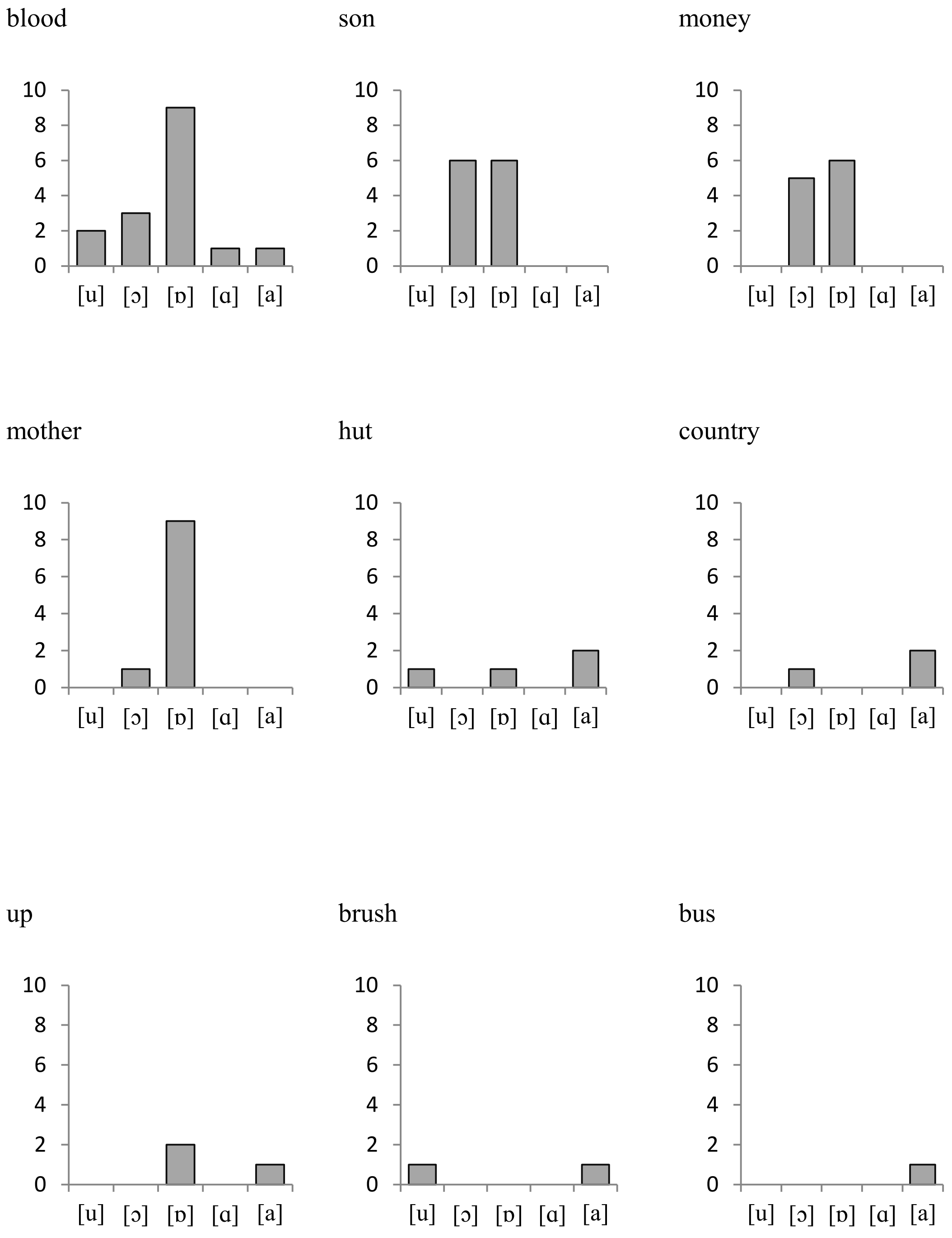

Figure 6: The Polish vowel system (Jassem 2003:105)

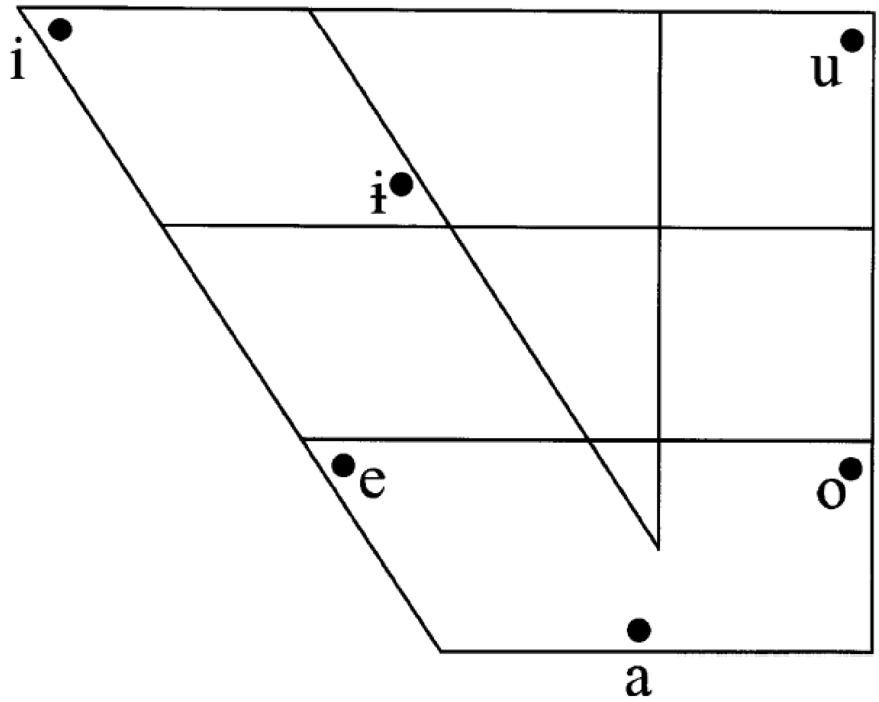

Figure 7: Chart showing the mean auditory STRUT value for each speaker in both tasks.