



1-1-1996

Variation in /ai/ in Northern British English, with comments on Canadian Raising

James Milroy

Variation in /ai/ in Northern British English, with comments on Canadian Raising

Variation in /ai/ in Northern British English, with comments on Canadian Raising

James Milroy
University of Michigan and University of Newcastle upon Tyne

1 Introduction

The diphthong /ai/ exhibits a good deal of variation in English generally with a clear allophonic split in some dialects between what has been called a 'raised' variant before voiceless obstruents and a phonetically quite distinct low-nucleus variant in certain other environments. The best known examples of this split are Canadian English and Scottish English, which exhibit, respectively, 'Canadian Raising' and the 'Scottish Vowel Length Rule' (SVLR). The chief aim of this paper is to investigate variation in this diphthong in Newcastle upon Tyne using data collected from 32 speakers in a current sociolinguistic project. I will concentrate on variation in terms of class, gender and age of speaker and on the length and quality distinction in /ai/ predicted by the Scottish Vowel Length Rule (SVLR).

The SVLR has been much discussed recently, particularly by McMahon (1991, 1994). Following Lass (1976) it is also known as 'Aitken's Law', and according to Lass's statement it describes what is essentially a typological difference between the Scots and English vowel systems. While the latter has phonemic length (some phonemes are distinguished from others by a length difference), Scots does not have phonemic length. Instead Scots vowel length (SVL) is allophonic, and, apart from two (or possibly three) vowels which are always short, all vowels are long or short according to environment. Before voiced fricatives, /r/ and in final position, vowels are long; otherwise they are short. It is important, however, to distinguish Scots vowel length from the 'voicing effect' that is present in English generally, presumably including Scots (see further McMahon, 1994: 62-3). Whereas this voicing effect predicts that vowels are slightly longer before voiced than before voiceless consonants, the Scots vowel length environments have markedly longer vowels than other environments. The length difference is easily audible, and it can reasonably be estimated that the duration of the SVL vowels is approximately twice that of the short vowels. In the description of class, gender and age patterns in this paper, I will assume a two-way distinction between the SVLR environments and others, describing the other vowels as 'short'.

The purpose of this paper is to report the findings of a systematic quantitative study of the SVLR as it affects the diphthong /ai/. The data-base is tape-recordings of 32 speakers of two social classes in Newcastle upon Tyne from a current ESRC-supported project (my co-workers are Lesley Milroy, Gerry Docherty, Paul Foulkes, Penny Oxley and David Walshaw). The analysis and quantification have been carried out on the conversational styles of these speakers, which is of course more representative of their usage than word-list or questionnaire-based data or data obtained in laboratory settings. As much of our knowledge of the SVLR is derived from less detailed studies than this, it is to be hoped that an analysis of this kind will contribute to our knowledge of the rule and help in assessing the reliability of the accounts that have been given.

The findings are: 1) that the Scottish Vowel Length Rule, as it affects /ai/, can be said to be present in the Tyneside community, but that it is undergoing change; 2) that in one respect the rule may not have been correctly stated; and 3) that the main change in

progress affects the short variants of /ai/ rather than the long variants. We will proceed by sketching in some of the general background.

2 The Scottish Vowel Length Rule: background to the Tyneside data

Although the SVLR is described as 'Scottish', it does appear that it should be present in Northumberland and Tyneside. The reasons for this are partly historical and partly based on general observation of Tyneside speech. The traditional dialects of North-eastern England share with Lowland Scots an origin in Old Northumbrian, and Middle English texts from NE England have much in common with Scots texts. Furthermore, observation of Tyneside speakers has confirmed that for some speakers at least, items such as *feet*, *fight* have short vowels, as predicted by the SVLR. If the rule is present and if it is in a process of change, it is expected that the variable (ai) will show different sociolinguistic distributions in terms of speaker variables such as age, sex and social class.

The characteristic local variant of the short vowel in Newcastle is regionally marked and has a high-mid nucleus near [ei] rather than the mid (central) nucleus usually described for the SVLR. The geographical distribution of this high-mid vowel variant is not clear, but it is found across the Scottish border in south-west Scotland, and in Northern Ireland, where it is characteristic of Belfast vernacular speech. Data-analysis by J Milroy (1982) shows that the items *bide*, *bride*, *Friday*, *wide*, *tile*, *mile*, *file*, *while*, *rise*, *drive*, occur with [ei] in Galloway, South-west Scotland. The geographical distribution, in so far as it can be determined, suggests that it is a relic form that may at one time have had a wider distribution, and this view may be supported by arguing that this vowel represents an early stage of the diphthongisation of Middle English /i:/ in the Great Vowel Shift. Other similarities between the Newcastle and Belfast vowel systems, in particular the incidence of ingliding diphthongs where southern British dialects have upgliding diphthongs, tends to strengthen the probability that we are dealing with quite conservative vowel systems in these locations. Unfortunately, Survey of English Dialects (SED) data (Orton, Sanderson and Widdowson, 1978) are of limited value in describing the rural background to the (ai) pattern, as many of the lexical items selected appear with unshifted [i:]. Moreover, Southern Scotland is not covered, and, presumably, the SED lexical items were not selected with the SVLR in mind. Even so, a number of relevant items have [ei] in Northumberland, Durham and North Cumberland, and the final vowel items *sky* and *thigh* alternate in Northumberland between [a:i] (as predicted by the SVLR), and [ei]. On this alternation, see further below.

The (ai) variable, as we have analysed and quantified it, is expected to reflect the allophonic split between long and short vowels, with particular lexical items favouring one or the other. As is well known, the SVLR results in alternations of various kinds, with, for example, *knife* (short) having a mid-vowel nucleus and *knives* (long) a low nucleus. In some dialects (e.g. Ulster Scots) the difference may be phonemic as it is reported that a series of words are differentiated as low or mid. These include *lie* ('recline': low) and *lie* ('tell lies': mid), *mine* (poss pron: low) and *mine* (e.g., coal-mine: mid) (Gregg 1964: 173). As implied above, /ai/ is clearly differentiated by the height of the nucleus, the long variant having a low nucleus and the short variants having a mid-central to rather high-front nucleus. It is usually assumed that this qualitative difference co-varies 100% with the length difference, i.e., that the length difference predicts the qualitative difference. According to the SVLR therefore, the open nucleus (as in [a:i]) should be confined to the

long environments (as in *size*, *sigh*), with the central to high-front nucleus occurring otherwise before voiced and voiceless consonants (e.g., in *side*, *sight*).

It should be noted that this differentiates the rule from Canadian Raising of /ai/ (Chambers 1973), which is phonetically and structurally similar. The CR raised variants are similar to the Scottish short variants in items such as *might*, *mice*. Indeed, this allophone is also widely heard in the northern states of the U.S.A. The low-nucleus variant in north America is also qualitatively similar to the Scottish long variant. The difference between CR and the SVLR lies in the distribution of the variants. Whereas the SVLR predicts the short raised variant before voiced consonants excluding voiced fricatives and /r/, CR reportedly has [a:i] before all voiced consonants, including items such as *side*, *dine*. The following data from a young speaker from Saginaw, Michigan, suggest that the phenomenon is variable just as the application of the SVLR is variable.

(1) 'CANADIAN' RAISING OF /ai/ (S.E. MICHIGAN)

| | |
|--------------|----------|
| short/raised | long/low |
| FIGHT | SIDE |
| KNIFE | SIZE |
| KNIVES | SIGN |
| CIDER | SIGH |
| DINER | SIGHED |
| FRIDAY | FIVE |
| FIRE | JULY |

Forms like *diner* seem to be variable, but *cider* usually has the short vowel where the long vowel is expected (as in *side*). The short vowel (with a centering glide) is found here before -[r] (and this seems very general in the region), and the forms of 'knife' do not show the alternation predicted by the consonant change. The SVLR is also usually said to have lexical exceptions, and it has sometimes been noted that while certain items occur categorically with [a:i] in long environments, others with similar following environments exhibit variation between, e.g., [ei] and [a:i] (cf. the variation in *sky*, *thigh* noted above).

The Tyneside situation is further complicated by the fact that many English dialects to the south have [ai] in all environments, with little or no length and height difference. Current work on the Midland city of Derby, for example, within our present research project, strongly suggests that a vowel near [a:i] with a lengthened nucleus occurs in all relevant environments - before voiceless and voiced consonants and finally. Conservative northern dialects may exhibit a lexical split between [a:i] and [ei], with monophthongisation to [a:] and to [i:] (as in 'neet' for 'night') occurring sometimes. Monophthongisation to [i:] is noted sporadically in the lexical items *night*, *right*, amongst older male speakers in the Tyneside sample, but it is very rare in the data collected. It is, however, the possible northward spread of the low-nucleus [ai] variants that may initially be expected to affect the SVLR short variants in Tyneside, transferring some lexical items, such as *wide*, *wine*, from the high-mid class to the low nucleus class. We shall assess this possibility below.

In a variationist study that is accountable to a large amount of data, it is not to be expected that the SVLR will be categorically observed by speakers, and it had been noted prior to quantification that some items in which [a:i] is predicted, such as *die* might occur with [ei], or a similar vowel, and that this is normally short. Similarly (and this is supported by word-list data), whereas *five* seems normally to have [a:i], *knives* normally occurs with [ei], failing to exhibit the SVLR alternation with *knife* (as in Ulster Scots according to Gregg (1964) and in the north American variety in Table 1, above). In the present project, three relevant studies have been carried out on (ai). These are: 1) quantification of 1114 variants used by 32 speakers in terms of age, gender and social

class; 2) instrumental measurement of the word-list items *sight*, *side*, *size* for 25 speakers in terms of vowel duration; 3) quantification of what appear to be 'violations' of the rule in terms of age, gender and social class.

3 The sociolinguistic study

It was decided to recognise three variants in terms of vowel quality rather than length. These were 1) low nucleus (approximately [a:i]); 2) central nucleus ([ɪ]); and 3) high nucleus (approximately [eɪ]). It was known that many lexical items might not show the whole range of variation; for example, *sight* might prove to vary between 2 and 3, but not 1, and *size* might vary between 1 and 2, but not 3. Indeed, some lexical items might not appear to vary in these terms at all, but be categorically in one realisation. It was nevertheless decided to quantify over both sub-categories of words (the [eɪ] words and the [aɪ] words) on the grounds that the number of tokens is great enough to ensure a reasonably balanced representation of both types of word. In the event, the results show very highly significant differences in preference for the different realisations of this variable in terms of class, gender and age of speaker, with the high-mid-nucleus localised variant [eɪ] being more favoured by working class speakers, males and (contrary to what might be expected) younger speakers. The figures for each group are presented in (2), below, in terms of social class, starting with working class speakers.

| | | | | |
|-----|---------------|---------|-------|--------|
| (2) | Variants | 1 [a:i] | 2 [ɪ] | 3 [eɪ] |
| | WORKING CLASS | | | |
| | Older F | 44 | 43 | 63 |
| | Older M | 26 | 18 | 95 |
| | Younger F | 28 | 26 | 81 |
| | Younger M | 29 | 18 | 92 |

The mid and high-mid variants (2 and 3) are of primary interest here. Males clearly use the localized 'vernacular' variant ([eɪ]) more than females do (column 3, above), and the scores for males are similar in both generations. Amongst the females, it is the younger women who favour the vernacular variant most. Thus, there is no support here for any theory that might propose that there is a change in progress towards the [a:i] variant, either in the male or the female scores. Thus, working class speakers in Tyneside are apparently not being strongly influenced by more southerly dialects of English or by RP, in which a variant similar to [a:i] is normally found throughout the range. Later in this paper, we further consider the use of the [a:i] variant, which is highest in the older female group. In (3) below, we consider the middle class scores:

| | | | | |
|-----|--------------|---------|-------|--------|
| (3) | Variants | 1 [a:i] | 2 [ɪ] | 3 [eɪ] |
| | MIDDLE CLASS | | | |
| | Older F | 51 | 72 | 16 |
| | Older M | 46 | 36 | 51 |
| | Younger F | 27 | 80 | 31 |
| | Younger M | 29 | 47 | 65 |

If we compare these in general with the working-class preferences, the main finding is that there is some movement away from the raised variant [eɪ] towards the mid nucleus variant, but that the gender differentiation is similar. The scores for variant 3 ([eɪ]) are in every case

lower than for the WC groups, but the males again score higher for [ei] than the females. The older males, however, have lower scores for this variant than the younger males and higher scores than the younger males for [a:i]. Again, this does not support the argument that there is a change in progress towards [a:i]. The high female scores for variant 2 (central nucleus) are strongly affected by the avoidance of the (localised) high nucleus variant in words of the type *light, like, type*, i.e., pre-voiceless stop environments. Males in both social classes show much less inclination to avoid the high-nucleus localised variant, but younger female speakers nevertheless favour this localised variant more than the older females. For younger speakers the scores for variant 1 ([a:i]) are about the same in both social classes, and the differences between younger speakers in terms of social class are carried in the different distributions of variants 2 and 3. Although a slight trend towards the low nucleus is present in the older generation, the most important differences are in the mid and high mid variants. The lexical distribution of items is exemplified in (4), below: items in column 1 are the items that the SVLR would predict as long, and items in columns 2 and 3 conform mainly to the SVLR short class as usually described. The table shows the pattern for an older male speaker:

(4) LEXICAL PATTERN FOR THE SVLR

| Variant: | 1 (low) | 2 (mid) | 3 (high-mid) |
|----------|-----------|----------|------------------|
| | five | right | mind |
| | quietened | Crikey | outside |
| | July | Byker | night |
| | died | Carlisle | bikes |
| | driving | | miles |
| | driving | | bikes |
| | | | cycling |
| | | | nine |
| | | | United |
| | | | times |
| | | | my |
| | | | (10 other items) |

This conforms closely to what the SVLR predicts, except that the item *my* has a high-mid short vowel rather than the long low vowel predicted by the rule. We consider such cases further below. First, we consider briefly the main results of statistical tests that have been carried out on the class x gender x age results. The use of log-linear modelling shows the effect of social class as overwhelmingly significant ($p < < < 0.001$) and that of gender as only slightly less so ($p < < < 0.001$). This means that the class and age differences are so overwhelmingly great that it is virtually inconceivable that they could have occurred by chance. The class effect is shown in (5), below (numbers are shown first, with percentages in parentheses):

(5) CLASS EFFECT FOR (ai)

| Variants | 1 | 2 | 3 | Total |
|----------|------------|------------|------------|-------|
| MC | 153 (27.8) | 235 (42.6) | 163 (29.6) | 551 |
| WC | 127 (22.5) | 105 (18.7) | 331 (58.8) | 563 |

The significant result is due to the very large differences between the classes in columns 2 and 3, and the difference in use of the low nucleus is slight (27.8 - 22.5%). Middle class

speakers strongly prefer the mid nucleus and working class speakers the high-mid nucleus, in words of the type *sight*, *side*, *sign*. The preferences are very marked: MC speakers use variant 2 more than twice as often as WC speakers; WC speakers use variant 3 ([ei]) about twice as often as MC speakers.

The effect of gender is slightly less significant but still very clear. The figures are given in (6):

(6) GENDER EFFECT FOR (ai)

| Variants | 1 | 2 | 3 | Total |
|----------|------------|------------|------------|-------|
| Female | 150 (26.7) | 221 (39.3) | 191 (34.0) | 562 |
| Male | 130 (23.6) | 119 (21.6) | 303 (54.9) | 552 |

Males have a much higher score than females for the vernacular localised variant 3 ([ei]), whereas females have a much higher score than males for the mid nucleus variant 2. Notice, however, that for both class and gender, the differences in variant 1 ([a:i]) are quite small, suggesting that there is little movement in preference for this realisation.

The effect of age is given in (7), below:

(7) AGE EFFECT FOR (ai)

| Variants | 1 | 2 | 3 | Total |
|----------|------------|------------|------------|-------|
| Older | 167 (29.8) | 169 (30.1) | 225 (40.1) | 561 |
| Younger | 113 (20.4) | 171 (30.9) | 269 (48.6) | 553 |

The contrast here comes, not in the intermediate variant 2 (scores for this are very similar, although they are different in terms of class and gender), but in variants 1 and 3, with younger speakers low on variant 1 and clearly favouring the localised vernacular variant ([ei]). The age distribution could be taken to suggest that there is a change in progress towards the vernacular variants, although the scores may be affected by the tendency among adolescents and young adults to favour the localised vernacular. It certainly does not support any argument that there may be a change in progress towards the southern and RP low nucleus variant. On the contrary, the figures may be interpreted as evidence of a swing towards affirmation of local identity in this variable (as in Martha's Vineyard (Labov 1972) and Belfast (a) (Milroy and Milroy 1985)). If there is such a swing it is seemingly led by males, as is now expected for localised changes.

4 Vowel duration measurements

The results displayed in Tables 1 - 3, above, all support the conclusion that the SVLR pattern is present for (ai) in some form. In order to pursue this point further, the word-list items *sight*, *side*, *size* and *sigh* were subjected to duration measurement by Jimmy Harnsberger at the University of Michigan. These represent the word-list (not conversational style) pronunciations of 25 speakers on those recordings that were suitable for instrumental analysis. We report here on the duration of the last three items for each speaker calculated as a ratio of the duration of these items in comparison with the item SIGHT, which for each speaker is calculated as 1.00. Thus the scores for the three items represent the percentage by which their duration is longer than that for SIGHT. The average of the 25 scores is shown in (8), below:

(8) DURATION RATIO (AVERAGE)

SIDE: 1.36.
 SIZE: 1.57
 SIGH: 1.57

The vowel in the item SIDE is on average 36% longer than that in SIGHT. The vowel in SIZE and SIGH is on average 57% longer than that in SIGHT. This confirms that the vowel is normally longer before voiced stops than before voiceless and further that it is still longer before voiced fricatives and finally. This is sufficient to show that the SVLR is present in the community for this vowel, and individual scores suggest that it is present in the majority of speakers, but not present in others. One older middle class female, for example, has a score of 1.40 for SIDE (40% longer than for SIGHT); for SIZE, however, it is 2.20 and for SIGH 2.30. This speaker certainly has the rule, as the SVLR items are more than twice as long as SIGHT and 57-64 percent longer than for SIDE. The distinction is, however, much less clear in some younger speakers, and this is why the average scores in (7), above, are not more dramatically differentiated. The scores for W (male) and V (female), both younger middle class, are given in (9):

(9) DURATION RATIOS FOR TWO YOUNGER SPEAKERS

| | W | V |
|------|------|------|
| SIDE | 1.97 | 1.25 |
| SIZE | 2.14 | 1.27 |
| SIGH | 1.69 | 1.30 |

If these trends prove to be more general, there is some reason to suspect that the SVLR is being gradually lost in the community. It must, however, be recalled that the duration measurements are based on word-list style and on vowel-duration rather than on vowel-height, and that SIGH may sometimes occur with a high-mid vowel and thus be a short vowel for some speakers. These considerations may account for patterns like that of W, above. In the next section, we consider some apparent violations of the SVLR that occur in the data.

6 'Violations' of the Scottish Vowel Length Rule

If the vowel length rule (as normally stated) is violated, then we will expect items in which the short vowel is predicted (such as wide) to occur with the long low-nucleus vowel, and items in which the long vowel is predicted (such as sigh, surprise) to occur with the short vowel with a high-mid or mid nucleus. A count of these exceptions is given in Table 10, over.

(10) 'VIOLATIONS' OF THE SVLR IN TYNESIDE

| | low for raised V | raised for low V |
|-------------------|------------------|------------------|
| MIDDLE CLASS: | | |
| Females (older) | 5 | 0 |
| Males (older) | 4 | 1 |
| Females (younger) | 2 | 14 |
| Males (younger) | 0 | 7 |
| TOTAL (MC) | 11 | 22 |
| WORKING CLASS: | | |
| Females (older) | 2 | 10 |
| Males (older) | 0 | 5 |
| Females (younger) | 1 | 14 |
| Males (younger) | 0 | 7 |
| TOTAL (WC) | 3 | 36 |

Clearly, there are many fewer apparent violations of the rule in the direction of 'low for raised' than of 'raised for low', and these are more common in the middle class than in the working class. Examples are: *pride, wine, final, nine, side, site, tidy, cycle*. This seems to suggest that some speakers may be aware that the low vowel carries some kind of supra-local prestige, and the pattern as a whole certainly fits Labov's idea that we are dealing in such cases with a sporadic correction strategy that is not at this stage involved in any vigorous pattern of language change. The 'raised for low' tokens, however, are much more numerous, and the social class pattern is the converse: i.e., 'violations' are more numerous in the working class than in the middle class. Examples of the items involved are: (mid for low) *tied, higher, Brian, drive, July, sky, buy, try identify* (final vowel), *my* (repeatedly); (high-mid for low) *knives, hibernating, Irish, dioxide, Micra* (a car model), *my* (repeatedly). There are strong reasons to suppose that these exceptions are not violations of the SVLR at all, but that many belong to a genuine alternating set, which predicts alternation in open syllable and other environments. The high-mid or mid vowel is well attested in Ulster Scots (Gregg 1964) and rural Galloway Scots (Milroy 1982) in words of this kind. For example Gregg (1964: 174) notes that the plural of 'knife' does not have the low vowel as predicted by the SVLR 'knife/ knives' alternation (seemingly *drive, rise, thrive, strive* do not normally have it either). In fact the SVLR alternation can certainly fail in Scottish speakers also (recent observation of a working class Glasgow speaker suggests this strongly). It is the open syllable items, however, that seem to have the short high-mid or mid vowel most often in Tyneside, and, as we have noted above, there is some SED atlas evidence that these are alternants in hinterland dialects. The occurrence of a number of the high-mid alternants in the Tyneside data does not constitute evidence that there is a change in progress transferring low vowel items to the mid or high-mid class; on the contrary, many such items must have belonged to the high-mid class in the past; they are now alternating, and the general trend may well be one of very gradual transfer from a traditional high-mid vowel to a low nucleus vowel.

7 General conclusions

The sociolinguistic analysis of (ai) in terms of class, gender and age does not suggest that there is a change in progress towards the low nucleus vowel found in more southerly dialects of English, but that the SVL distinction between two sets of lexical items is being

maintained in the community with some indication of change in progress towards the more localised variants. However, apparent violations of the SVLR in favour of the high-mid and mid vowel realisations for items of the type *buy, try* should not be interpreted as showing a change in progress away from the SVLR low vowel values. On the contrary, it appears that in some respects the SVLR as usually stated is not a correct representation of speaker usage of /ai/ in some relevant varieties. Speakers do not universally realise items like *try, diary, knives* with a long low-nucleus diphthong either in Newcastle or in Scotland and Northern Ireland. It appears from comparative evidence that the high-mid vowel is the conservative vowel in the SVLR environments, and the Tyneside results may be interpreted as evidence that these environments are alternating environments, in which the change by transfer to the low vowel is being resisted. Table 2, above, shows that younger speakers are particularly likely to use high-mid/ mid vowels in these environments; therefore, there is no evidence of a generational movement away from the high-mid/ mid realisations. Here, as elsewhere, more localised realizations are being maintained. Any change in progress that is discernible in the data depends on female, and to some extent, middle class preferences for the mid vowel over the high-mid vowel [ei]. This, if it is present, is however, very gradual, and it appears that the localized [ei] in items of the type *right, wide* is in no danger of dying out.

References

- Chambers, Jack K. (1973). "Canadian Raising." *Canadian Journal of Linguistics* 18:113-35.
- Gregg, R. J. (1964). "Scotch-Irish urban speech in Ulster," in *Ulster dialects: an introductory symposium* (Holywood, Co. Down: Ulster Folk Museum) 163-92.
- Labov, William (1972). *Sociolinguistic patterns*. (Philadelphia: University of Pennsylvania Press).
- Lass, R. (1976). *English phonology and phonological theory*. (Cambridge: Cambridge University Press).
- McMahon, April M. S. (1991). "Lexical phonology and sound change: the case of the Scottish vowel length rule" *Journal of Linguistics* 27: 29-53.
- McMahon, April M. S. (1994). *Understanding language change*. (Cambridge: Cambridge University Press).
- Milroy, James (1982). "Some connections between Galloway and Ulster speech." *Scottish Language* 1:23-27.
- Milroy, James and Milroy, Lesley (1985). "Linguistic change, social network and speaker innovation." *Journal of Linguistics* 21:339-84.
- Orton, H., Sanderson, S and Widdowson, J. D. A. (1978). *The Linguistic Atlas of England*. (London: Croom Helm).