The Social Perception of a Sound Change in York, Northern England

Daniel Lawrence

1 Introduction

A core finding of sociolinguistics is that sound changes may become available as a social-semiotic resource as they propagate through a speech community. The social evaluations attached to linguistic innovations are often used as an explanation for differences in the trajectories of parallel changes (e.g. Labov et al. 2014), or to explain differences in the rates of adoption of innovations across social groups (e.g. Hall-Lew, 2009). While much work has framed these evaluations by referring to their *salience* or the degree of *stigma* attached to them, recent studies have attempted to develop a more sophisticated understanding of the types of social meaning communities may attach to variation. In particular, the notion of the *persona* has risen to prominence – it is argued that speaker-listeners' social perceptions of linguistic variation tend to be structured around locally-relevant stereotypical figures such as the 'valley girl' or 'burnout' (Eckert, 2008; Moore & Podesva, 2009; D'onofrio, 2015). Applying this perspective to sound change, it has been argued that innovations which come about due to global processes of change (such as chain-shifting principles) may become attached to local stereotypes in certain speech communities, resulting in divergent behavior on the part of some groups of speakers.

While this approach provides an intuitive account of how global processes of change might interact with local social-semiotic systems in influencing the speech patterns of a given community, it raises the issue of how the social meanings proposed can be verified, and how their role in constraining the change under study can be diagnosed. This paper approaches this problem by combining the acoustic analysis of a sound change in progress with the results of an innovative social perception experiment which tested listeners' social intuitions regarding phonetic variation in the changing forms. This strategy enables the quantification of the degree to which listeners associate different speech forms with different social stereotypes, allowing quantitative predictions to be formed regarding which forms are associated with which social meanings, and for which groups of speaker-listeners.

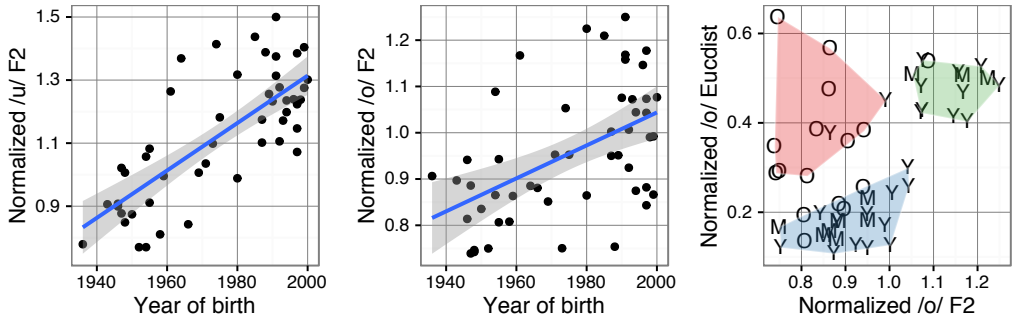
2 /u/ and /o/ fronting in York

The data for the present study are taken from a corpus of recordings collected from a convenience sample of 52 York residents born between 1935 and 2000. F1 and F2 measurements were taken at 20 equidistant points along the vowel trajectory, and these values were normalized using the normalization method of Watt & Fabricius (2002). Figures 1(a) and 1(b) plot these measurements as a function of speaker year of birth. Both vowels show evidence of fronting, with the most rapid and incremental fronting affecting /u/. The fronting of /o/ appears to be less regular, in the sense that speakers are less tightly clustered around the regression line in 1(b) than 1(a).

Plotting Euclidean distance and F2 values in two-dimensional space demonstrate how change in /o/ involves both fronting and diphthongization. This is demonstrated in Figure 1(c) The colored hulls show the output of a density-based clustering algorithm on these data, which identifies three groups of speakers – those with back /o/ realizations and middle/high Euclidean distances; those with back /o/ realizations and low Euclidean distances, and those with fronted /o/ realizations and high Euclidean distances. It should also be noted that membership of each cluster is predictive of the age of the speaker – back, diphthongal speakers are most likely to be in the oldest age group; back, monophthongal speakers are most likely to be in the middle or youngest age group, and speakers who produce front diphthongs are most likely to be in the youngest age group. The pattern appears to involve a reduction of the range of variation in diphthongization of back /o/ variants, and the adoption of very fronted, diphthongal variants among a subgroup of younger speakers.

In summary, the production data provide evidence for the fronting of /u/ and /o/ in this community. While /u/ fronting appears to proceed in a relatively uniform manner, /o/ fronting appears to

Figure 1: Variation and change in /o/ and /u/



(a) (b) (c)

have taken place more slowly and spread less uniformly across the population. Considering the role of /o/ diphthongization in this process reveals that /o/ fronting is restricted to speakers with diphthongal /o/ – while the fronting of monophthongal /o/ is in principle possible, and even attested in other Northern British varieties of English (Watt & Tillotson, 2001), there appears to be an absence of fronted, monophthongal speakers in this sample, evident in the lack of datapoints in the bottom right-hand quadrant of 1(c).

These data are generally consistent with Haddican et al.'s (2013) recent findings on /u/ and /o/ fronting in York. The patterning of these changes raises two key questions:

i) What causes /u/ to front so rapidly and uniformly in comparison to /o/?

ii) Why do younger speakers avoid fronting monophthongal /o/?

While a number of social, cognitive and linguistic factors might be invoked to explain these patterns, the remainder of this paper will focus on the proposal of Haddican et al. (2013), whose explanation centers around the social indexicality of variation in the changing vowels. With regard to the relative uniformity of change in /u/ vis-a-vis /o/, the authors propose that the fronting of /u/ is less important as a social-indexical resource to York speakers in comparison to the fronting and diphthongization of /o/. Under this account, this relative lack of social indexing allows /u/ fronting to propagate rapidly and relatively uniformly across the community, since the adoption of fronted /u/ variants will have limited social consequence for speakers. In contrast, /o/ diphthongization is claimed to be strongly associated with socioeconomic status and 'local' regional identity. Such claims are reasonable, since the diphthongization of the mid vowels is known to be a shibboleth of Northern/Southern English identity, at least in production (Watt, 2000; Beal, 2004). The implication of Haddican et al.'s (2013) proposal is that change in /o/ might be slower and less uniform than change in /u/ due to the participation of /o/ in a pattern of social indexing.

The author's third key claim is that fronted, monophthongal /o/ is associated with a stigmatized working-class stereotype, the 'chav'. This stereotype, reflecting an individual who is typically unemployed and engages in antisocial behaviour, has become a key feature of popular discourses around social class in the United Kingdom (see e.g. Hayward & Yar, 2006). Haddican et al. (2013) provide convincing evidence from ethnographic interviews that this social category is important to younger York residents, and propose that the absence of fronted, monophthongal /o/ in their data may be explained by younger speakers' avoidance of forms enregistered as 'chav' speech.

Haddican et al.'s (2013) social-indexical account of variation and change in /o/ and /u/ represents a common pattern of argument in sociolinguistic work – inferring a social signalling system based on observed production patterns. While their explanation is intuitive, and convincingly supported by their ethnographic data, it should be recognised that it represents one possibility in a potentially very large explanation space. In order to evaluate this proposal, the present work attempts to formulate and test a set of concrete predictions regarding the social perception of variation

in the target vowels. One way of thinking about any claim regarding the social meaning of phonetic variation is to express it as model of social categories distributed in multidimensional phonetic space. These social categories may include 'macro-level' meanings such as LOCAL or WORKING CLASS, or more locally-specific personae such as CHAV. An example for /u/ and /o/ variation in York, based on Haddican et al. (2013), is provided in Figure 2, which represents the categories LOCAL WORKING CLASS and CHAV in F2-Euclidean distance space.

Figure 2: Hypothesised mappings of indexical meanings to F2-Euclidean distance space

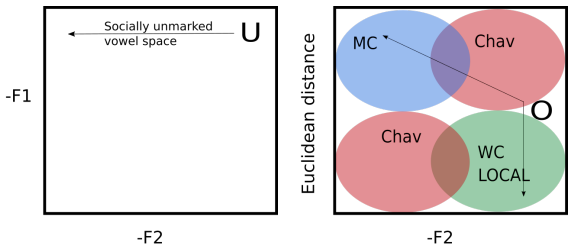


Figure 2 includes an addition to the argument presented in Haddican et al. (2013), with the inclusion of the hypothesis that back /o/ monophthongs are also enregistered as part of the CHAV persona. Since back /o/ diphthongs as well as fronted /o/ monophthongs are absent from the speech of younger speakers, it is reasonable to extend Haddican et al.'s (2013) proposal to include these forms.

Understanding a claim regarding social indexicality in the manner shown in Figure 2 allows concrete predictions regarding the social perception of the linguistic variables under study. For example, when a listener hears a talker use a back, diphthongal variant of /o/, we predict that they are most likely to assign that talker to the CHAV category. With increasing fronting, we predict that they will be increasingly more likely to assign MIDDLE CLASS to that talker. This pattern of reasoning can be extended across all combinations of phonetic parameters and social categories (see section x.x)

In order to test these predictions, a social perception experiment was conducted with the same speaker sample described in section x.x. The aim of this experiment was to quantify the extent to which York listeners associate the proposed social categories with /u/ and /o/ variation. The design of this experiment is described in the following section.

3 The social perception of /u/ and /o/ variation

3.1 Experimental design

3.1.1 Auditory stimuli

The auditory stimuli consisted of resynthesized tokens of the words *too*, *food* (/u/), *toast* and *so* (/o/), read by a 24-year-old middle-class speaker from York. The stimuli were resynthesized using a *Praat* script based on Alku et al.'s (1999) *IAIF* method. The complete set of /o/ stimuli included tokens representing three steps of fronting of monophthongal variants and three steps of two types of fronting (targeting the vowel onset and offglide) across diphthongal tokens. These included a back diphthongal realization, two steps of fronting at the vowel onset, and two steps of fronting of the offglide. The /U/ stimuli included examples of three levels of fronting, from a back realization to very fronted, as well as three identical tokens with lowered onsets, resulting in more diphthongal tokens.

Figures 3.1-3.3 give the IPA symbols which will be used to refer to the tokens in the analyses.

Figure 3.2: GOAT variants used in the experiment

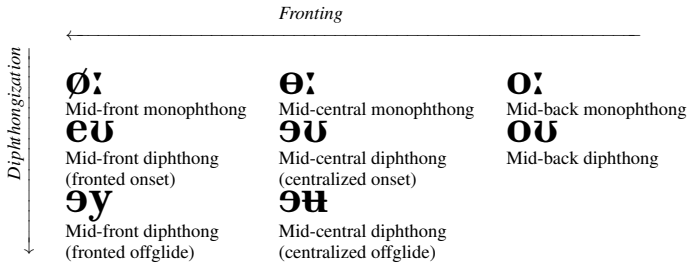
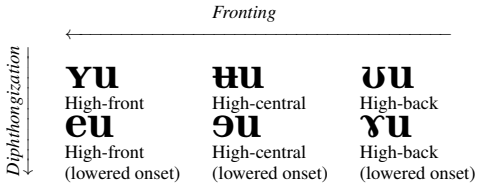


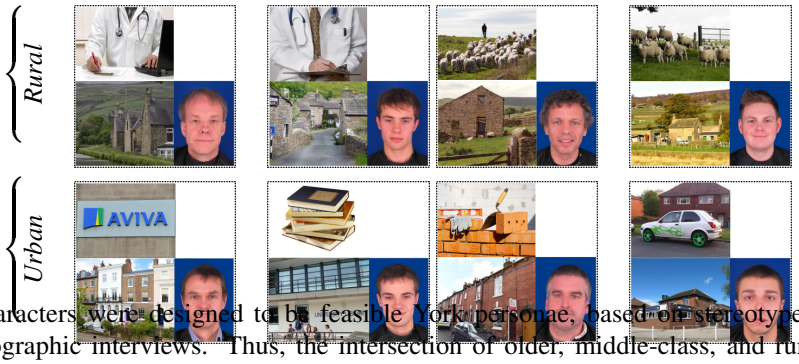
Figure 3.3: GOOSE variants used in the experiment



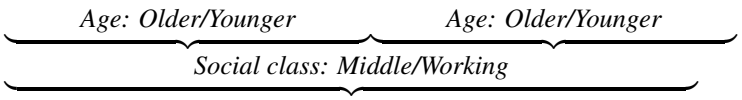
3.1.2 Visual stimuli

Visual stimuli consisted of eight images representing individuals of different ages (older vs younger), different occupations (working class vs middle class) and of different regional background (urban vs rural). The social dimensions of older/younger, working-class/middle-class, and urban/rural were included based on the findings of a set of ethnographic interviews conducted prior to the main data collection phase, as well as the claims made by Haddican et al. (2013). Each stimulus image contained three components – an image of a face (providing information about the character's age), an image of a place of work/study (providing information about the character's social status) and an image of an urban or rural location (providing information about the regional background of the character). The eight images represent all possible combinations of these three social dimensions, shown in Figure 3.7.

Figure 3: Visual stimuli



The characters were designed to be feasible York personae, based on stereotypes discussed in the ethnographic interviews. Thus, the intersection of older, middle-class, and rural is represented by a doctor in a rural Yorkshire village (a), while the corresponding middle-class character is represented as a middle-aged businessman associated with a well-known insurance company (e). In all cases, the choice of component images reflects comments made by participants during the ethnographic interviews. A key claim made by Haddican et. al. (2013) is that fronted monophthongal GOAT variants are associated with a working-class youth stereotype called a 'chav'. In this experiment, the 'chav' is treated as the intersection of the urban, young and working-class dimensions, represented by the character in image (h). The non-facial images were taken from



public domain collections, and the faces were selected from the Stirling ESRC facial database (http://pics.stir.ac.uk/ESRC/). The facial images were selected based on a pre-rating task completed by 20 students at the University of York.

3.1.3 Experimental task

Participants were told that they were listening to an actor preparing for a role in a sitcom set in York. In the training phase of the experiment, they were asked to categorize the visual stimuli based on a series of prompts (e.g. 'Which character is older', 'Which character is from rural Yorkshire?'). They achieved this with an accuracy of 96%. During main the experiment, participants saw two images per trial, heard a speech token, and were asked to select the character which they thought the actor was pretending to be. The two images on each trial differed in terms of one of the three social dimensions, with the remaining two kept constant between each image pair. For example, participants would see older and younger rural, working class characters in a single trial, but an older working class and younger middle class character would never appear together. Altogether, this resulted in 12 image combinations which were presented with each of the 32 speech samples. Participants were given two breaks at one-third and two-thirds of the experiment, where they were encouraged to take a brief rest and re-start when ready. The following section will outline the key findings from the production analysis, before using them to formulate and test predictions regarding the social perception data.

3.1.4 Predictions based on Haddican et al. (2013)

The perception data allow the explicit testing of these hypotheses – by analysing the way listeners assign variation in the changing forms to the social categories presented in the experiment, it is possible to verify the general pattern proposed above, as well as to explore the consistency with which individuals can identify the proposed social meanings. This is achieved through framing the hypotheses as statements about probabilities, then estimating those probabilities through a statistical model. Specifically:

(a) The probability of a WORKING-CLASS selection should be higher for monophthongal /o/ variants than diphthongal /o/ variants

(b) The probability of a LOCAL selection should be higher for monophthongal /o/ variants than diphthongal /o/ variants

(c) Variation in /u/ should have a comparatively smaller effect on the probability of a WORKING-CLASS or LOCAL selection compared to variation in /o/

(d) The probability of a CHAV selection should be higher for fronted, monophthongal /o/ than back, monophthongal /o/

(e) The probability of a CHAV selection should be higher for back, diphthongal /o/ than front, diphthongal /o/

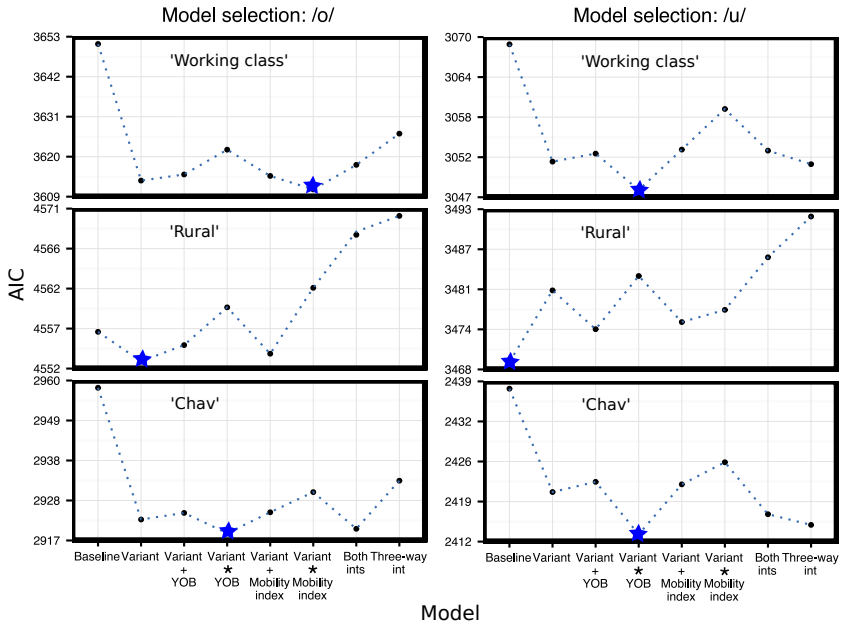
(f) (d) and (e) will interact with listener year of birth, with younger listeners showing a larger effect

3.2 Model selection

These probabilities were estimated using multilevel logistic regression fit for each vowel on each social dimension. The analyses presented here focus on models for the WORKING CLASS dimension

(all WORKING CLASS vs all MIDDLE CLASS images), the RURAL dimension (all RURAL vs all NON-RURAL images), and the CHAV category (the CHAV image vs all others). Baseline models predicted the selection of the social category modeled with random intercepts, by-listener random slopes for *variant*, and random slopes for sound sample within *variant*. The relative fit of these models was compared with more complex ones, including terms for the variant heard, the listener's year of birth, their score on the mobility index, and the interactions of those factors.

Figure 4: AIC for each candidate model, ordered by complexity



These models are compared in Figure 4 which plots the Akaike Information Criterion value for each model specification, arranged in order of model complexity. A lower AIC value indicates a better fit to the data. The best-fitting models are marked with a star. In terms of social class selections, the models for both /u/ and /o/ include the *variant* term, confirming that listeners were able to use variation in these vowels as a reliable cue to socioeconomic status and to the CHAV subcategory. RURALselections were reliably affected by variation in /o/, but not /u/. The best models for WORKING CLASS also include interaction terms for listener mobility (/o/) and year of birth (/u/), and the best models for CHAV include interaction terms for year of birth, indicating that listeners of different ages and social backgrounds behaved differently in the social perception task. Hypothesis tests were then conducted through a bootstrapped likelihood ratio test comparing models with and without each term (table x).

/o/	Chisq	Df	P(ChiSq)	/u/	Chisq	Df	P(ChiSq)
WORKING CLASS				**WORKING CLASS**			
Variant	51.64	7.00	<0.0001	*Variant*	27.29	5.00	<0.0001
*Variant*Mobility*	18.42	8.00	0.02	*Variant*Age*	16.18	6.00	0.01
RURAL				**CHAV**			
Variant	17.37	7.00	0.01	*Variant*	27.74	5.00	<0.0001
CHAV				*Variant*Age*	19.25	6.00	<0.01
Variant	49.39	7.00	<0.0001				
*Variant*Age*	19.68	8.00	0.01				

3.3 Results: Main effects

3.3.1 /u/

Figure 5: Main effects for /u/

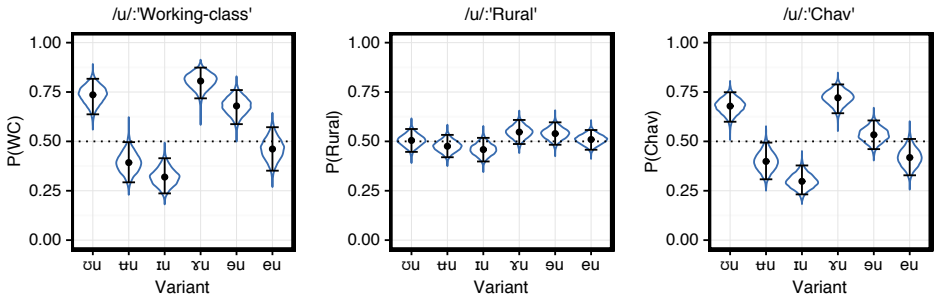


Figure 5 demonstrates how variation in /u/ affected listeners selections in terms of the social categories WORKING CLASS, RURAL and CHAV. The points represent the mean predicted probability of each selection, with error bars showing the upper and lower bounds of the Highest Posterior Density interval, estimated through 10000 posterior samples. There is a clear effect of fronting on WORKING-CLASS selections, whereby fronter /u/ variants are more likely to cue middle-class selections and back variants are more likely to cue working-class selections. /u/ diphthongization seems to be weakly associated with WORKING CLASS, in that diphthongization weakens the effect of fronting (evident in the shallower slope of fronting across more diphthongal /u/ variants, and in the fact that the HPD interval for fronted, diphthongal /u/ crosses the p=.5 line. /u/ variants show no reliable effect on listeners' RURAL selections, while the results for CHAV mirror those for WORKING CLASS. These results contrast with the predictions formed earlier, where it was suggested that /U/ variation would be perceived as relatively socially unmarked given its lack of social stratification in production. Despite the fact that /u/ fronting appears to have been adopted in a relatively uniform manner in this community, listeners consistently map innovative variants to the MIDDLE CLASS category.

3.3.2 /o/

Figure 6: Main effects for /o/

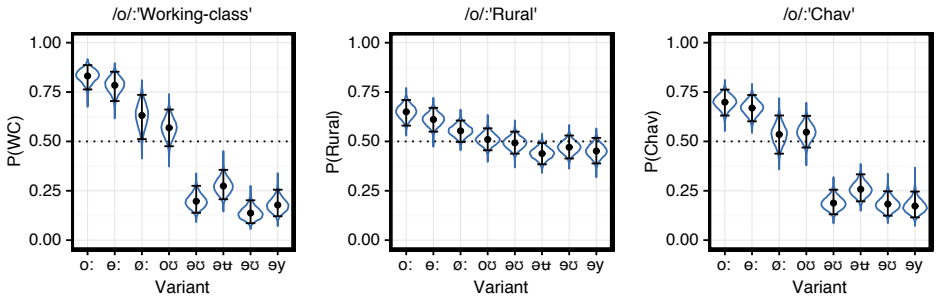


Figure 6 visualizes the effect variant on social selections for /o/. In terms of WORKING CLASS selections, there is a clear effect of diphthongization, with more diphthongal variants reliably cueing the selection of a middle class image, and monophthongal variants cueing the selection of working-class image. The only exception to this pattern is back, diphthongal /o/; the fact that the interval for this variant crosses zero means that we cannot be certain that it had a reliable effect in either

direction, but the fact that the interval is biased above the p=.5 line suggests that listeners tended to select working-class images when hearing this variant. Monophthongs all reliably predict the selection of a working-class image, but the strength of this effect weakens for more fronted variants. Note that this weakening effect of fronting holds for both diphthongs and monophthongs, in contrast to what might be predicted based on Haddican et al. (2013).

In terms of RURAL selections, /o/ shows a weak but reliable effect for monophthongal variants. While diphthongs all show no reliable effect (in that the HPD interval crosses p=.5), monophthongs reliably cue the selection of a rural character. The fact that only monophthongs show reliable effects suggests that while /o/ monophthongs are associated with rural speech, the converse is not true of diphthongs. Overall, the results are broadly consistent with our earlier predictions – listeners reliably map /o/ monophthongs to the RURAL category.

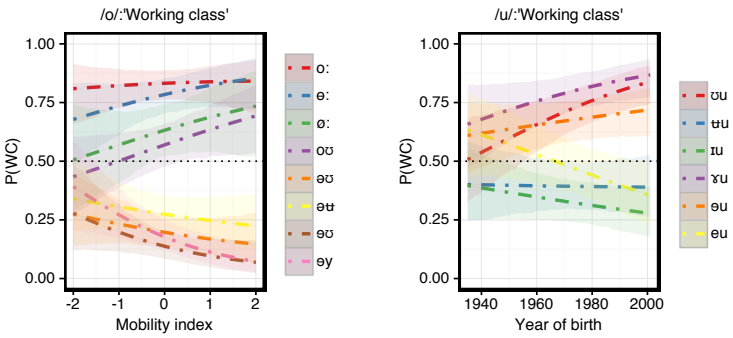
A key prediction based on Haddican et al.'s (2013) proposal was that fronted /o/ monophthongs would be associated with the CHAV subcategory. The data for /o/ provide no support for this proposal – rather, CHAV selections mirror the effects for general WORKING CLASS selections. While the most back /o/ variants reliably cue CHAV selections, fronted /o/ shows no reliable effect, and the overall trend is for fronting to *lower* the probablity of a CHAV selection – the opposite to what would be expected under the Haddican et al. (2013) account.

3.4 Results: Listener variation

A further prediction made in section 3.1.4 was that listeners would vary in their social perception of variation in /o/ and /u/. In particular, it was predicted that the CHAV association between fronted /o/ monophthongs and back /o/ diphthongs would be strongest among younger listeners. While no evidence was found for the predicted relationship between /o/ variation and CHAV, the data still provide evidence of variation in social perception, demonstrated below.

3.4.1 WORKING CLASS selections

Figure 7: Interaction effects: WORKING CLASS selections

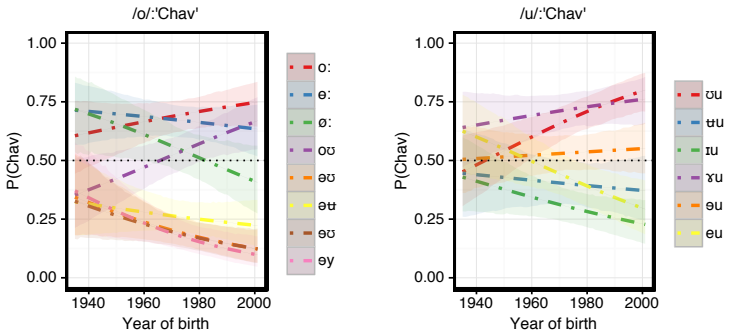


(a) (b)

Figure 7 demonstrates variation in social class selections as a function of listener year of birth and the mobility index. For /o/, there is a significant effect of listener mobility (7a), with more mobile listeners generally more sensitive to the monophthong-diphthong distinction as a cue to socioeconomic status. The largest variation appears to be in the perception of the most fronted diphthong and most back monophthong – these variants have a limited effect on the responses of less mobile listeners, suggesting that those listeners hear them as comparatively unmarked. In contrast, more mobile listeners tend to associated these variants with middle-class characters. /u/ selections show no evidence of variation in terms of mobility, but a significant effect of listener age (7b). Older listeners are generally less sensitive to fronting as a cue to socioeconomic status.

3.4.2 CHAV selections

Figure 8: Interaction effects: CHAV selections



(a) (b)

Figure 8 demonstrates the interaction of listener year of birth and variant heard on CHAV selections. In the case of /o/ (8a), it appears that younger listeners are more sensitive to diphthongs as an index of 'non-Chav', whilst back, diphthongal and fronted, monophthongal /o/ show an interesting 'cross-over' pattern: for older speakers, [oʊ] weakly disfavours CHAV selections, while for younger speakers it is reliably associated with CHAV; while [o:] reliably cues a CHAV selection for older listeners, it weakly disfavours such selections among younger listeners. For /u/ (8b), the effect of age is similar to that for WORKING-CLASS selections – younger listeners map the back-front dimension of /u/ more reliably to the CHAV character than older listeners.

4 Conclusion

References

[Beal(2004)] Beal, Joan. 2004. English dialects in the north of england: phonology. In *A Handbook of Varieties of English, Volume 1*, ed. B Kortmann and E W Schneider, volume 1, 122–144. Mouton.

[D'Onofrio(2015)] D'Onofrio, Annette. 2015. Persona-based information shapes linguistic perception: Valley girls and california vowels. *Journal of Sociolinguistics* 19:241–256.

[Eckert(2008)] Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12:453–476.

[Ester et al.(1996)Ester, Kriegel, Sander, and Xu] Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd* 96:226–231.

[Fruehwald(2013)] Fruehwald, Josef. 2013. The Phonological Influence on Phonetic Change. Doctoral dissertation, University of Pennsylvania.

[Haddican et al.(2013)Haddican, Foulkes, Hughes, and Richards] Haddican, Bill, Paul Foulkes, Vincent Hughes, and Hazel Richards. 2013. Interaction of social and linguistic constraints on two vowel changes in Northern England. *Language Variation and Change* 25:371–403.

[Hall-Lew(2009)] Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in a San Francisco neighborhood. Doctoral dissertation, Stanford University.

[Harrington et al.(2008)Harrington, Kleber, and Reubold] Harrington, Jonathan, Felicitas Kleber, and Ulrich Reubold. 2008. Compensation for coarticulation,/u/-fronting, and sound change in standard southern british: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123:2825–2835.

[Labov et al.(2013)Labov, Rosenfelder, and Fruehwald] Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89:30–65.

[Moore and Podesva(2009)] Moore, Emma, and Robert Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38:447–485.

[Watt and Fabricius(2002)] Watt, Dominic, and Anne Fabricius. 2002. Evaluation of a technique for improving the mapping of multiple speakers vowel spaces in the f1˜ f2 plane. *Leeds Working Papers in Linguistics and Phonetics* 9:159–173.

[Watt and Tillotson(2001)] Watt, Dominic, and Jennifer Tillotson. 2001. A spectrographic analysis of vowel fronting in bradford english. *English World-Wide* 22:269–303.

[Watt(2000)] Watt, Dominic JL. 2000. Phonetic parallels between the close-mid vowels of Tyneside English: Are they internally or externally motivated? *Language Variation and Change* 12:69–101.