

Sound change, priming, salience

Producing and perceiving variation
in Liverpool English

Marten Juskan

Draft
of 8th October 2018, 23:53

Language Variation

Editors: John Nerbonne, Dirk Geeraerts

In this series:

1. Côté, Marie-Hélène, Remco Knooihuizen and John Nerbonne (eds.). The future of dialects.
2. Schäfer, Lea. Sprachliche Imitation: Jiddisch in der deutschsprachigen Literatur (18.–20. Jahrhundert).
3. Juskan, Marten. Sound change, priming, salience: Producing and perceiving variation in Liverpool English.

ISSN: 2366-7818

Sound change, priming, salience

Producing and perceiving variation
in Liverpool English

Marten Juskan

Marten Juskan. 2018. *Sound change, priming, salience: Producing and perceiving variation in Liverpool English* (Language Variation 3). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/210>

© 2018, Marten Juskan

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/>

ISBN: 978-3-96110-119-1 (Digital)

978-3-96110-120-7 (Hardcover)

ISSN: 2366-7818

DOI:[10.5281/zenodo.1451308](https://doi.org/10.5281/zenodo.1451308)

Source code available from www.github.com/langsci/210

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=210

Cover and concept of design: Ulrike Harbort

Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}T_EX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

[./langsci/graphics/storage/logo.pdf](#)

To Daniela

Contents

1	Introduction	1
1.1	Intentions – what this study is about	1
1.2	Restrictions – what this study is not about	3
1.3	Structure of the book	4
2	A brief history of Liverpool and Scouse	7
2.1	The first 600 years	7
2.2	19 th century	10
2.3	20 th century	14
2.3.1	Enregisterment and the ‘Scouse industry’	14
2.3.2	Decline	16
2.3.3	Regeneration	19
2.4	21 st century – outlook	21
2.5	Summary	23
3	Variables	25
3.1	General remarks	25
3.2	Supragsegmentals	26
3.3	Consonants	27
3.3.1	/ŋ(g)/	28
3.3.2	Lenition (of /k/)	29
3.4	Vowels	32
3.4.1	happy	33
3.4.2	NURSE – SQUARE	35
3.5	Summary	38
4	A few words on salience and exemplar theory	39
4.1	Salience	39
4.1.1	Salience and circularity	39
4.1.2	Cognitive vs. social salience	41
4.1.3	Salience in this study	43

Contents

4.2	Exemplar theory	47
4.2.1	Basic principles	48
4.2.2	Application in (socio-)linguistics	49
4.2.3	Frequency and salience in exemplar theory	53
4.3	Summary	56
5	Interview method	57
5.1	Interview structure	57
5.2	Participants	59
5.3	Transcription	62
5.4	Measuring	62
5.4.1	Consonants	62
5.4.2	Vowels	65
5.4.3	Normalisation	66
5.5	Phonological context	74
5.6	Statistical analysis	76
6	Vowel production	81
6.1	happy	82
6.1.1	F1 (happy)	82
6.1.2	F2 (happy)	93
6.1.3	Synthesis and Pillai scores (happy)	104
6.2	NURSE	108
6.2.1	F1 (NURSE)	108
6.2.2	F2 (NURSE)	118
6.2.3	Synthesis and Pillai scores (NURSE)	131
7	Consonant production	141
7.1	/ŋ(g)/	142
7.1.1	Overview	142
7.1.2	Phonological context	142
7.1.3	Style and gender	145
7.1.4	Age and social class	147
7.1.5	Style shifting	150
7.2	/k/	151
7.2.1	Overview	151
7.2.2	Phonological context	153
7.2.3	Style and gender	155
7.2.4	Style and social class	157

7.2.5	Age and gender	159
7.2.6	Age and social class	162
7.2.7	Social class and gender	164
7.2.8	Style shifting	165
8	Awareness, comments, evaluation	171
8.1	Scouse and ‘Liverpoolness’	171
8.1.1	Accent and identity	171
8.1.2	Distinctness, geographical spread, and ‘plastic’ Scousers	172
8.1.3	In the north, but not of it?	174
8.2	Features of Scouse	177
8.2.1	Geographical variation	177
8.2.2	Suprasegmentals	178
8.2.3	Phonological variables	179
8.3	Evolution and evaluation of Scouse	185
8.3.1	Old speakers	185
8.3.2	Middle-aged speakers	188
8.3.3	Young speakers	191
9	Discussion (production)	197
9.1	happy: Indicator (of northernness)	197
9.1.1	Overall age differences	197
9.1.2	Gender and class	198
9.1.3	Style shifting and awareness	200
9.1.4	Classification	201
9.2	NURSE: Marker to stereotype and back again	201
9.2.1	Overall age differences	201
9.2.2	Gender and social class	204
9.2.3	Style shifting and awareness	206
9.2.4	Classification	208
9.3	Velar nasal plus: Indicator with prestige option	208
9.3.1	Age, class, and gender	208
9.3.2	Style shifting and awareness	210
9.3.3	Classification	211
9.4	Lenition: From indicator to stereotype	212
9.4.1	Age	212
9.4.2	Gender and class	213
9.4.3	Style shifting	214
9.4.4	Awareness and attitudes	216

Contents

9.4.5	Classification	217
9.5	Summary	218
10	Experiment design	221
10.1	Stimuli	221
10.1.1	Stimuli sentences and frequency of keywords	221
10.1.2	Vowel stimuli	225
10.1.3	Consonant stimuli	226
10.2	Presentation	228
10.2.1	Online platform	228
10.2.2	Reaction times	231
10.3	Participants	232
10.4	Statistical analysis	235
11	Perception results	237
11.1	happy	238
11.1.1	Overview	238
11.1.2	Prime	239
11.1.3	Position of carrier word	240
11.1.4	Geographical distance	241
11.1.5	Frequency	242
11.1.6	Age	244
11.2	NURSE	245
11.2.1	Overview	245
11.2.2	Prime	245
11.2.3	Position of carrier word	247
11.2.4	Stimulus order	248
11.3	/ŋ(g)/	251
11.3.1	Overview	251
11.3.2	Prime	252
11.3.3	Phonological context	253
11.3.4	Age	255
11.4	/k/	256
11.4.1	Overview	256
11.4.2	Prime and social class	257
11.4.3	Phonological context	261
11.4.4	Geographical distance	263
11.4.5	Frequency	265

12 Discussion (Perception)	267
12.1 Priming and salience	267
12.1.1 Salient vs. non-salient variables	267
12.1.2 Degree of priming and accuracy	268
12.1.3 Salience as likelihood of remembrance	269
12.2 Social factors	272
12.2.1 Social class	272
12.2.2 Gender	273
12.2.3 Age	274
12.2.4 Geographical distance	276
12.3 Non-social factors	277
12.3.1 Time held in memory	277
12.3.2 Stimulus order	279
12.3.3 Phonological environment	280
12.4 Issues and limitations	283
12.4.1 The problem of velar nasal plus	283
12.4.2 Comparability with previous research	285
12.4.3 The issue of frequency	286
12.4.4 Size of the priming effect	287
12.5 Direction of priming	287
12.5.1 The problem of replicating priming experiments	287
12.5.2 Hay et al.'s explanation for inverted effects	289
12.5.3 Assimilation and contrast effects	290
12.6 Summary and implications	293
13 Conclusion	295
Appendix A: Questionnaire	299
Appendix B: Reading passage	303
Appendix C: Word list	305
Appendix D: Stimuli for perception test	307
References	309

6 Vowel production

This chapter is based on the 20 participants of the primary sample who provided 7950 data points.¹ 4565 of these are observations of happy, 1770 concern NURSE, and 882 SQUARE (the remainder are instances of TRAP and FLEECE). The majority of the data were collected during the spontaneous part of the interview (n = 5245), while the word list (n = 761), the reading passage (n = 1361), and accent performance (n = 583) account for much smaller proportions. Table 6.1 provides an overview of how many realisations of the test vowels were made in which sub-sample of speakers. Within each age group, the sample seems to be fairly balanced with respect to gender and social class. There are some exceptions (e.g. the female-male ratio among middle-aged middle-class subjects), but this is only to be expected in a dataset that consists largely of spontaneous speech, and on the whole it does not seem as if a particular sub-group dominated the sample too much in numeric terms.

Table 6.1: Vowel observations by age, gender, and social class

		old		middle		young	
		f	m	f	m	f	m
happy	mc	307	201	411	705	335	362
	wc	197	248	506	380	354	559
NURSE	mc	145	78	179	334	107	129
	wc	87	98	152	157	155	149
SQUARE	mc	75	48	92	94	65	89
	wc	50	55	63	77	90	84

The two dependent variables F1 and F2 are analysed separately, both because it is well possible that change manifests itself with respect to one dimension but not the other (cf. Harrington 2006), and because this means they can be more easily

¹Early versions of this and the next chapter were presented at the 21st LIPP-Symposium in Munich and at an internal workshop at the University of Freiburg. I am very grateful to the audiences for the valuable feedback received on these occasions. All remaining shortcomings are entirely mine.

compared to the two consonantal variables /k/ and velar nasal plus, for which there are only one-dimensional measurements (cf. §5.4.1). §6.1.3 and §6.2.3 briefly look at the two dimensions in conjunction, largely on the basis of F1-F2 vowel plots which are familiar from many works on phonetics and phonology. Contrary to the sections that deal with F1 and F2 individually, the focus in §6.1.3 and §6.2.3 will be much less on social predictors of vowel realisation other than age, and more on a summary synthesis of the style shifting results and a – very general – contextualisation of happy and NURSE with FLEECE and SQUARE, respectively.

All of the figures in this chapter (except those in §6.1.3 and §6.2.3) are designed in such a way that higher values on the y-axis represent ‘more Scouse’ variants. For NURSE, this is not straightforward: both raising (towards [e]) or lowering (towards [ɛ]) could be interpreted as ‘more Scouse’ as a range of different front vowels are cited in the literature as possible local realisations. In this book, [ɛ] was chosen as it is often considered to be the most typical target of the NURSE-SQUARE merger in Liverpool. When present, p-values in the boxplots (rounded to three decimal places) correspond to t-tests comparing the oldest to the middle-aged, and the middle-aged to the youngest group. The value that is horizontally centred is always taken from the t-test comparing means in the oldest and the youngest group. Arithmetic means are marked by dots in box plots, whereas bold horizontal lines in the middle of the notches indicate the relevant median. The extent of these notches, in turn, roughly corresponds to the 95% confidence intervals of the medians. Whiskers extend to the most extreme data points that are still within 1.5 times the interquartile range. All plots were generated with ‘ggplot2’ (Wickham 2009) and dynamically integrated into this book with ‘knitr’ (Xie 2015). As mentioned in §5.4.3, all vowel plots and analyses are based on the Watt-Fabricius normalised F1 and F2 values (coded as “F1W” and “F2W” respectively). Whenever tests on the “raw data” (as opposed to the estimates of the regression models) are reported, it is equally these normalised F1 and F2 values that are referred to, and *not* the original, absolute measurements in Hz (which would be the ‘raw data’ in the narrow sense).

6.1 happy

6.1.1 F1 (happy)

6.1.1.1 Overview

The maximal model for happy F1 measurements as described in §5.6 suffered from severe collinearity ($\kappa = 40.27$). Much of this collinearity turned out to be unproblematic, though, because it held between interactions and its constituent main ef-

fects (a model without these interactions had $\kappa = 30.62$) and could therefore safely be ignored. Furthermore, separate regressions of normalised F1 values on manner and place of preceding ($\kappa = 95.67$), and following sound ($\kappa = 18.53$) showed that these factors were (moderately) collinear, too, so only manner was retained from each pair. The maximal model without the predictors place of articulation of preceding and following sound (and also, for the sake of diagnostics, without any interactions) then only exhibited a moderate and acceptable degree of collinearity ($\kappa = 15.09$). Interactions were then re-entered into the regression and the minimal adequate model shown below was arrived at based on AIC scores and F-tests comparing nested models. The minimal adequate model for F1 of happy (R^2 -equivalent = 0.178) is reprinted as Table 6.2.

Table 6.2: happy (F1): mixed linear effects regression

Fixed effects:	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	0.69	0.03	178.02	24.42	< 0.001	***
STYLElist	-0.06	0.03	1901.45	-2.04	0.04	*
STYLEread	0.02	0.01	1553.02	1.33	0.18	
STYLEfree	0.05	0.01	954.42	4.72	< 0.001	***
AGE56-85	-0.01	0.00	2114.86	-2.96	< 0.01	**
AGE30-55	-0.00	0.00	2109.14	-0.33	0.74	
GENDERf	-0.02	0.00	2111.94	-5.87	< 0.001	***
CLASSmc	0.01	0.00	2115.13	4.06	< 0.001	***
ZIPF	0.01	0.01	106.00	1.73	0.09	.
DURATION	-0.00	0.00	2114.64	-2.53	0.01	*
POSTMANNERaffr	-0.05	0.02	2103.57	-3.05	< 0.01	**
POSTMANNERfric	0.01	0.01	2103.23	1.00	0.32	
POSTMANNERgli	0.00	0.01	2076.64	0.80	0.42	
POSTMANNERliq	0.01	0.01	1993.46	1.11	0.27	
POSTMANNERnas	0.04	0.01	2105.78	5.05	< 0.001	***
AGE56-85:GENDERf	-0.03	0.00	2115.77	-6.71	< 0.001	***
AGE30-55:GENDERf	0.01	0.00	2114.16	3.14	< 0.01	**
GENDERf:CLASSmc	0.01	0.00	2093.34	2.69	0.01	**
Random effects:	(number of obs: 2116, groups: WORD, 221)					
Groups	Name	Variance	Std.Dev.			
WORD	(Intercept)	0.001	0.038			
Residual		0.015	0.122			

Style is a significant predictor, as are age, gender, and social class of participant. Gender furthermore significantly interacts with age and social class, respectively. The phonetic and phonological factors duration and (manner of) following sound also have a statistically significant effect. Frequency of carrier word was not sig-

nificant at the 5% threshold, but an ANOVA revealed that removing this effect from the model nevertheless resulted in a significantly worse fit to the data. Since frequency *does* qualify as a statistical trend it was therefore kept in the model. The two phonetic predictors will be briefly discussed first.

6.1.1.2 Phonological context

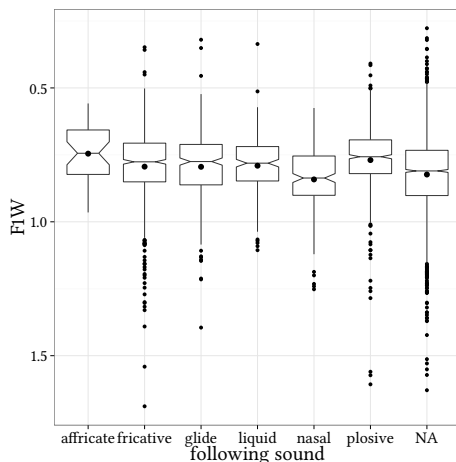


Figure 6.1: happy (F1) by following sound

Figure 6.1 shows normalised F1 values on the y-axis, and the manner of the consonant that follows the test vowel on the x-axis. Just as in the mixed-effects model, only two levels of this factor really stand out: when happy is followed by an affricate F1 values are slightly lower, i.e. the vowel is a bit higher, whereas when happy precedes a nasal, F1 is significantly higher, so the vowel is somewhat lower. All the other contexts have similar means (black dots) and medians (black horizontal bars around the middle of the boxes), and do not differ significantly from each other. It is unclear why happy should be higher when the following word (happy is, of course, always word-final) starts with an affricate, but it should be noted that the number of observations in this sub-category is rather small anyway ($n = 37$), so these figures might not be very representative. For nasals, on the other hand, this is less of an argument, as there are much more data in this sample ($n = 194$), but here the shift can potentially be explained on phonetic grounds. As an effect of regressive assimilation, it is likely that happy exhibits some degree of nasalisation when it occurs before a nasal, and nasalisation is known to shift F1 upwards (House & Stevens 1956), which explains the

difference visible in the graph. Interestingly, F1 seems to be very similar in cases when happy is not directly followed by anything because it is the last sound in a stretch of speech (this context is coded as “NA” in Figure 6.1), but it is a lot less straightforward why happy should be lower before a pause.

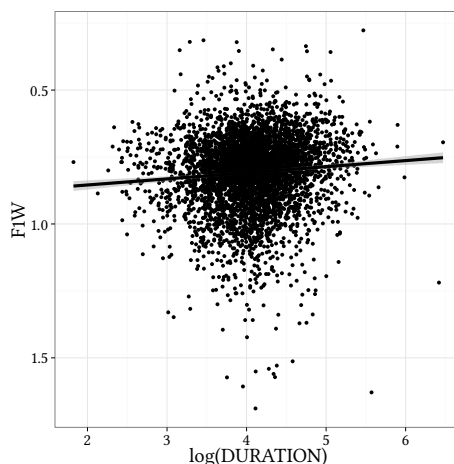


Figure 6.2: happy (F1) by duration

The effect of the other phonetic predictor, duration, seems to be more straightforward. For this fixed effect, a negative coefficient was returned in the mixed linear effects regression, which translates into the regression line we see in Figure 6.2, with duration on the x-axis (log-transformed for better visualisation) and estimated F1 on the y-axis. The y-axis is inverted (just as in the other F1 plots of happy) because higher F1 values, i.e. lower realisations, actually indicate less Liverpool-like variants of happy. The regression line has an upward slope, which means that longer happy realisations have lower F1 values. This is hardly surprising and does not call for a special explanation. Rather, it is a general phonetic principle that, all other things being equal, longer vowels will also tend to be more peripheral because there is more time for the tongue to reach its final position. Or, put another way, if a vowel is going to be short, there is *not enough* time for the tongue to move into a comparatively extreme position, so realisation of the vowel will be more central.

6.1.1.3 Frequency

Despite the fact that it is only a statistical trend we will also briefly look at the impact of frequency on F1 values in another regression plot. If one considers the

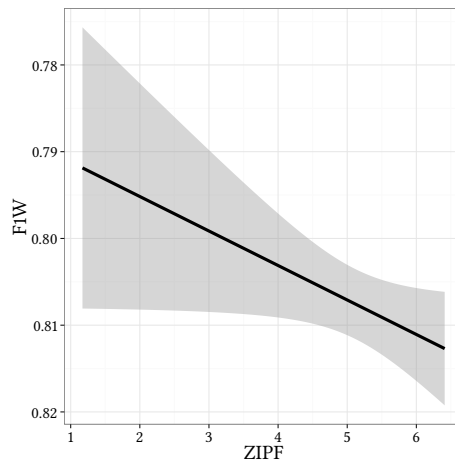


Figure 6.3: happy (F1) by frequency of carrier word

scales of the y-axes in Figure 6.2 and Figure 6.3 it becomes immediately obvious that both frequency and duration have only a rather small influence on F1. Higher Zipf scores on the x-axis indicate more frequent carrier words, so the bottom line of this graph is that the more frequent a word containing *happy* is, the lower the realisation of *happy* will be – which is something that is also implied by the positive correlation coefficient in the mixed-effects model. This is in line with what we know about frequency effects in general, viz. that higher frequencies of use favour phonetic reduction processes. It should be borne in mind, however, that my sample is based on spoken and relatively informal language, and therefore consists almost exclusively of high frequency words, anyway (Zipf scores for *happy* carrier words: mean = 5.114, first quantile = 4.57, median = 5.37, third quantile = 5.79). Data on low frequency words are comparatively scarce, which also shows in the larger standard deviation for these tokens (cf. the larger dark grey area on the left-hand side of Figure 6.3). It would be interesting to see if the tendency described above would also surface (possibly in a statistically more robust way) in a sample that is more balanced in this respect.

6.1.1.4 Class and gender

We will now turn towards the first social predictor, class of speaker. The mixed-effects regression model found a positive correlation coefficient for middle-class interviewees. This highly significant correlation indicates that middle-class speakers produced lower *happy* realisations (higher F1 values) than their working-class

counterparts. This in itself is interesting (and somewhat unexpected), but instead of going into this class difference by itself, I would like to immediately include the significant interaction of social class and gender that the mixed-effects model revealed as well. Said interaction is visualised in Figure 6.4a and Figure 6.4b.

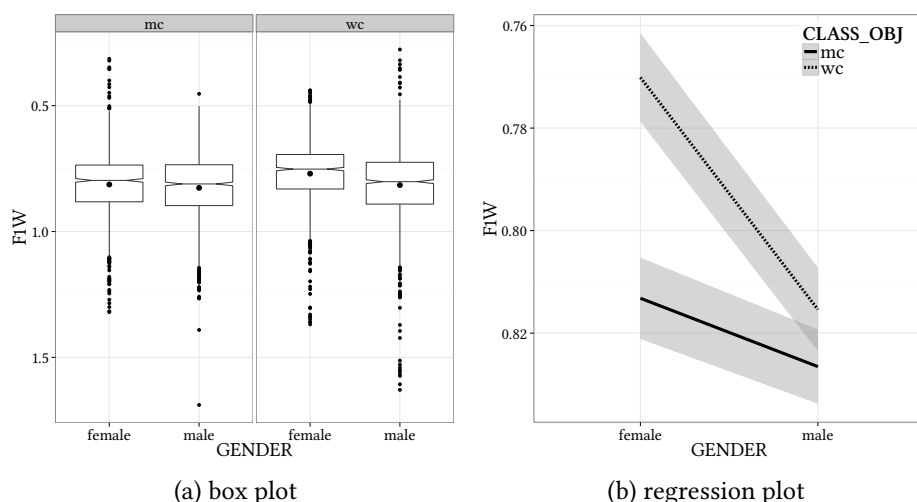


Figure 6.4: happy (F1) by gender and class

The regression plot on the right shows estimated F1 on the y-axis and gender of subject on the x-axis. Social class is coded by line type, with dashed representing working-class, and solid marking middle-class subjects. This plot shows very nicely that the class difference seems to be mostly driven by female speakers, because, for these subjects, the two lines are comparatively far apart (estimated F1 for middle-class women is just above 0.81, whereas the value for working-class females is about 0.77), and the standard deviations are clearly distinct from one another. For males, on the other hand, the estimates are relatively similar for both social classes, and, what is more, the standard errors are not distinct (which is visualised by the grey, partially overlapping areas on the right hand side of the plot). T-tests on the raw data confirm this description: the difference between middle and working class is highly significant for female ($t(2107.847) = 7.591$, $p < 0.001$), but (just about) insignificant for male speakers ($t(2342.064) = 1.92$, $p = 0.055$). Furthermore, the steep slope of the working class line as opposed to the comparatively flat one for middle-class subjects suggests that gender differences (which were also found to be a significant main effect in the mixed-effects regression) are mostly found in the working class.

The box plots in Figure 6.4a visualise the same gender X social class interaction as Figure 6.4b, but this time the focus is on the gender, instead of the class difference. In the left panel (middle class), the two boxes for female and male participants can be seen to occupy essentially the same space. The means (black dots) and medians (horizontal bars) are very similar, although those of the males do seem to be slightly lower, and the confidence intervals of the medians (as represented by the notches) also appear to overlap to a certain extent. A t-test finds that male and female realisations of happy *are*, in fact, significantly different in height for middle-class speakers ($t(2249.52) = -2.437$, $p = 0.015$), but – as outlined below – it is both less significant and less pronounced than in the other group. Among working-class speakers, on the other hand, men have clearly lower realisations than women, as can be seen in the right panel of the graph. Perhaps we should rather say, that working-class *women* have significantly *higher* realisations ($t(2235.848) = -7.549$, $p < 0.001$), because it seems to be this group in particular which behaves differently. Working-class men actually have F1 values which are relatively similar to those of both middle-class men and women.

6.1.1.5 Age and gender

Before analysing the other significant interaction gender was involved in, we will first look at age of speaker as a significant main effect. Remember that one of the hypotheses that this study intends to test is that “Scouse is getting Scouser”, i.e. that younger speakers have more extreme local variants and/or use these local variants more often than older speakers. Figure 6.5 plots the average F1 values of happy for the three groups of speakers (oldest subjects on the left, middle-aged group in the middle, youngest speakers on the right). The graph shows (a) that there is considerably less variation in the youngest group (indicated by the size of the box, and the extremeness of outliers), compared to the older speakers, and (b) that all three groups are significantly different from one another (cf. the p-values included in the figure).

Most importantly, however, it is obvious from Figure 6.5 that the hypothesis that younger speakers are, on average, more Scouse cannot be confirmed with respect to the height of happy. Both mean and (to a lesser extent) median values decrease ‘graphically’ (due to the inverted y-axis; in numeric terms they of course *increase*) from left to right. Average F1 thus significantly increases from the old to the middle-aged group ($t(1838.232) = 2.257$, $p = 0.024$), and from the middle-aged to the young ($t(3609.648) = 2.949$, $p = 0.003$), meaning that happy systematically becomes *lower*, i.e. less Scouse, the younger the speaker. Liverpoolians between the ages of 19 and 29 are therefore actually the *least* Scouse in the sample, because

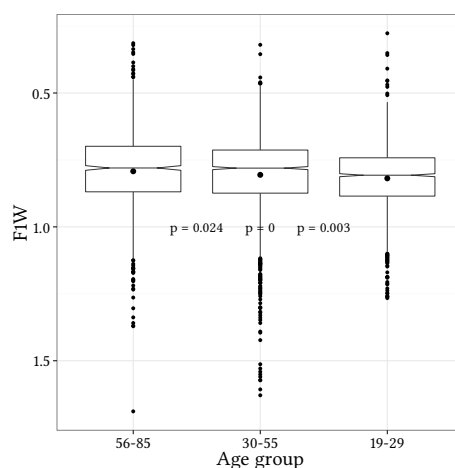


Figure 6.5: happy (F1) by age

they have lower realisations of happy than their parents' generation, who in turn use lower variants than the oldest speakers.

How does this predictor interact with gender, then? In Figure 6.6b, we again (cf. Figure 6.4b) see gender of participant on the x-axis and estimated F1 on the y-axis, but it is now age of participant instead of social class which is coded by line type (solid for the oldest, dotted for the middle-aged, and dashed for the youngest speakers). Just as in the corresponding graph for gender and social class, it is obvious that women seem to be behind the age differences described in the preceding paragraphs. Only for female subjects are the lines clearly distinct from one another. For male speakers, on the right-hand side, the regression lines of all age groups are either in, or at least very close to, the dark grey area which marks overlapping standard deviations. This indicates that age groups are not significantly different from one another if only the male subjects are considered. Table 6.3 shows that this is indeed the case: t-tests of all three pairings (comparing the old speakers to the middle-aged, the middle-aged to the young, and the young to the old) yield highly significant results for female, and non-significant ones for male subjects across the board.

Just as in the case of the gender X social class interaction, the differences in slope of the three regression lines (steep fall for the oldest, moderate fall for the middle-aged, and ever so slight rise in the youngest group) also suggest that, conversely, gender differences are not equally pronounced in the different age groups. Indeed, Figure 6.6a shows that there is a quite pronounced difference

Table 6.3: happy (F1): t-tests of age by gender

test	women			men		
	t	df	p	t	df	p
old-middle	4.062	1091.324	< 0.001	-1.368	798.429	0.172
middle-young	5.062	1596.832	< 0.001	-0.333	1999.537	0.739
young-old	8.689	888.424	< 0.001	-1.624	768.511	0.105

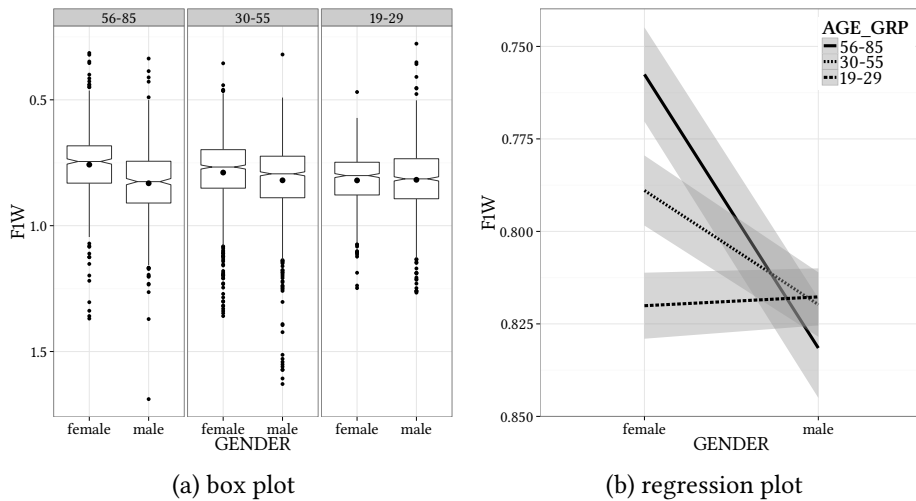


Figure 6.6: happy (F1) by gender and age

between men and women in the oldest group, with the former having significantly lower realisations than the women ($t(896.946) = -7.773$, $p < 0.001$). This difference is still present in the middle group ($t(1955.552) = -4.707$, $p < 0.001$), although the distance has clearly decreased. For the youngest speakers, finally, there is no longer a significant difference between male and female speakers ($t(1606.076) = 0.411$, $p = 0.681$). It is interesting that men do not seem to have changed much across these three generations; their means are comparatively similar. Women, on the other hand, appear to have adapted to the men over time by lowering their originally higher happy realisation to one that is almost identical to that of men.

6.1.1.6 Style shifting

The last factor that turned up as a significant fixed effect in the regression model is style, the variable which this study is most interested in (though closely followed by age) because style shifting is considered as an epiphenomenon of salience. Figure 6.7 represents the style dimension along the x-axis, starting with the word list on the left and going through the reading passage and free, spontaneous speech to the accent imitation task on the right of the graph. F1 is, as usual, marked on the y-axis, while line type once again codes the three age groups of speakers. The size of the whiskers corresponds to the standard error for each register and age group.

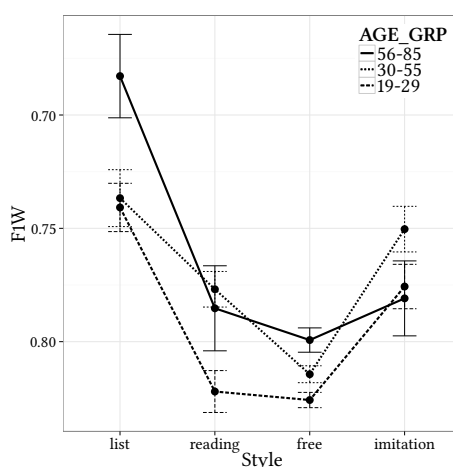


Figure 6.7: happy (F1) by age and style

The crucial question now is whether we are looking at style shifting or not. Productions of happy, across age groups, are higher in the ‘list’ register than when subjects read out a text. These differences are visibly significant because the whiskers of the ‘list’ and ‘reading’ categories do not overlap for any of the age groups. From the reading passage to spontaneous speech happy seems to get even lower, but this is only significant in the middle-aged group. Both the whiskers attached to the solid (old speakers) and the dashed (young speakers) dots overlap for these two registers. Realisations during the accent performance are then higher again, though not as high as when reading a word list (the rise is non-significant for the oldest speakers in the sample). Older speakers thus only distinguish the word list style from the other three (which do not differ significantly), middle-aged subjects have similar realisations for the word list

and the accent performance, and the youngest interviewees distinguish reading and spontaneous speech together from both imitation and the word list.

These slightly different tendencies in the three age groups are not pronounced enough to show up as a significant interaction of age group and style in the mixed linear effects regression. Judging from Figure 6.7, this is probably not too surprising because there does seem to be a similar trend across the age groups even if the differences are not all equally significant. The pattern that we see is not really one of ‘classical’ style shifting, though. Instead of a steady decline from the most formal to the most informal speech style, a sort of U- or V-shaped pattern emerges. In this context, it seems worthwhile to consider an explanation based on phonetic aspects, more precisely on duration. Table 6.4 reports mean durations of happy (in milliseconds), and mean frequency scores of carrier words for each style and in each age group.

Table 6.4: happy: durations (ms) and frequency (Zipf scores) by style and age

	list		reading		free		imitation	
	dur.	freq.	dur.	freq.	dur.	freq.	dur.	freq.
old	90.20	4.69	64.30	4.46	68.16	5.10	69.31	4.34
middle	90.82	4.72	62.70	4.46	62.94	5.23	62.13	4.51
young	104.07	4.69	62.72	4.41	65.36	5.37	55.94	4.43

As is to be expected, vowel durations are, on average, considerably longer when people read out a word list. Longer vowel durations, in turn, favour more peripheral happy realisations, as has been shown (and explained) above. At least in parts, the higher variants in the word list can thus be explained simply by the fact that they are also longer. It is clear, however, that this is only part of the story, and that some sub-conscious shifting must be involved as well, because vowel durations are comparable (and certainly not considerably longer) during accent imitation, text reading, and free speech. Higher, more Scouse realisations when performing the accent can therefore not be explained by a phonetic effect of longer vowel durations. The same seems to hold, more or less, for frequency of the carrier word as well: more frequent words, on average, are used in spontaneous speech (which could go some way to explaining lower realisations in this register), but in the other three styles frequencies are very similar, despite the fact that F1 values are not. A possible explanation that goes beyond duration and frequency will be discussed in Chapter 9.

6.1.2 F2 (happy)

6.1.2.1 Overview

Since the maximal model for happy F2 measurements was based on exactly the same dataset as the one where the dependent variable was F1, the same problems with collinearity also emerged. These were dealt with in an identical manner as has been described for the happy F1 model. In the end, interactions were likewise re-entered into the regression and model selection based on AIC scores and F-tests comparing nested models resulted in the minimal adequate model printed as Table 6.5 (R^2 -equivalent = 0.268).

Table 6.5: happy (F2): mixed linear effects regression

Fixed effects:	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	1.64	0.02	656.54	75.97	< 0.001	***
STYLElist	0.10	0.05	1799.79	1.98	0.05	*
STYLEread	-0.05	0.02	1638.38	-2.55	0.01	*
STYLEfree	-0.04	0.02	1501.28	-2.09	0.04	*
AGE56-85	0.03	0.02	1959.22	1.41	0.16	
AGE30-55	0.01	0.02	1947.73	0.36	0.72	
GENDERf	0.02	0.00	2079.61	6.38	< 0.001	***
CLASSmc	-0.01	0.00	2081.99	-1.88	0.06	.
DURATION	0.00	0.00	2070.57	7.80	< 0.001	***
PREMANNERaffr	0.00	0.04	115.37	0.01	1.00	
PREMANNERfric	0.05	0.02	125.52	2.19	0.03	*
PREMANNERliq	-0.04	0.01	114.24	-3.22	< 0.01	**
PREMANNERnas	-0.02	0.02	93.43	-0.92	0.36	
POSTMANNERaffr	0.06	0.02	2071.42	3.23	< 0.01	**
POSTMANNERfric	-0.01	0.01	2077.31	-2.22	0.03	*
POSTMANNERgli	0.00	0.01	2063.14	0.32	0.75	
POSTMANNERliq	-0.05	0.01	2025.37	-5.72	< 0.001	***
POSTMANNERnas	0.01	0.01	2067.51	0.69	0.49	
STYLElist:GENDERf	0.09	0.02	1982.28	5.36	< 0.001	***
STYLEread:GENDERf	-0.03	0.01	1962.17	-3.22	< 0.01	**
AGE56-85:CLASSmc	-0.01	0.00	2080.77	-2.93	< 0.01	**
AGE30-55:CLASSmc	0.01	0.00	2070.63	3.18	< 0.01	**
STYLEread:AGE56-85:GENDERf	-0.02	0.03	1950.17	-0.54	0.59	
STYLEfree:AGE56-85:GENDERf	-0.03	0.02	1988.95	-1.22	0.22	
STYLEimit:AGE56-85:GENDERf	-0.01	0.04	1945.32	-0.34	0.74	
STYLEread:AGE30-55:GENDERf	-0.01	0.03	1945.26	-0.22	0.82	

6 Vowel production

STYLEfree:AGE30-55:GENDERf	0.01	0.02	1965.59	0.37	0.71
STYLEimit:AGE30-55:GENDERf	0.05	0.03	1944.32	1.61	0.11
STYLElist:AGE56-85:GENDERm	-0.02	0.10	2063.72	-0.19	0.85
STYLEread:AGE56-85:GENDERm	0.00	0.03	1949.44	0.07	0.95
STYLEfree:AGE56-85:GENDERm	-0.04	0.02	1978.35	-1.75	0.08
STYLElist:AGE30-55:GENDERm	0.10	0.07	2081.56	1.39	0.17
STYLEread:AGE30-55:GENDERm	-0.00	0.03	1942.75	-0.08	0.94
STYLEfree:AGE30-55:GENDERm	0.04	0.02	1963.18	1.75	0.08
<hr/>					
Random effects:	(number of obs: 2116, groups: WORD, 221)				
Groups	Name	Variance	Std.Dev.		
WORD	(Intercept)	0.002	0.043		
Residual		0.015	0.124		

This minimal model is very similar to the one that was reported for F1 measurements. Style and gender turn up as significant predictors again. Age is not a significant main effect for F2 of happy, but it does appear in a significant interaction of age and class. Social class, in turn, just about fails to reach significance as a main effect at the 5% level. The second two-way interaction that was retained in the model is that of style and gender. A three-way interaction of style, age, and gender did not reach significance, but was retained anyway because an ANOVA revealed that eliminating it resulted in a significantly worse fit to the data. With respect to the non-social predictors there are some changes as well. Vowel duration is, once more, highly significant, but frequency of the keyword does not seem to have a statistically robust impact on F2 measurements. Contrary to the regression of F1, both the following *and* the preceding consonant (or, rather, its manner of articulation) are significant fixed effects in this model. These last two predictors will be briefly analysed first.

6.1.2.2 Phonological context

Figure 6.8a is a box plot of F2 values (on the y-axis) sorted by preceding (manner of) consonant (on the x-axis). “NA” here stands for observations where no phonemic transcription was available for the carrier word and which, as a consequence, could not be coded for preceding sound (this mostly concerned proper

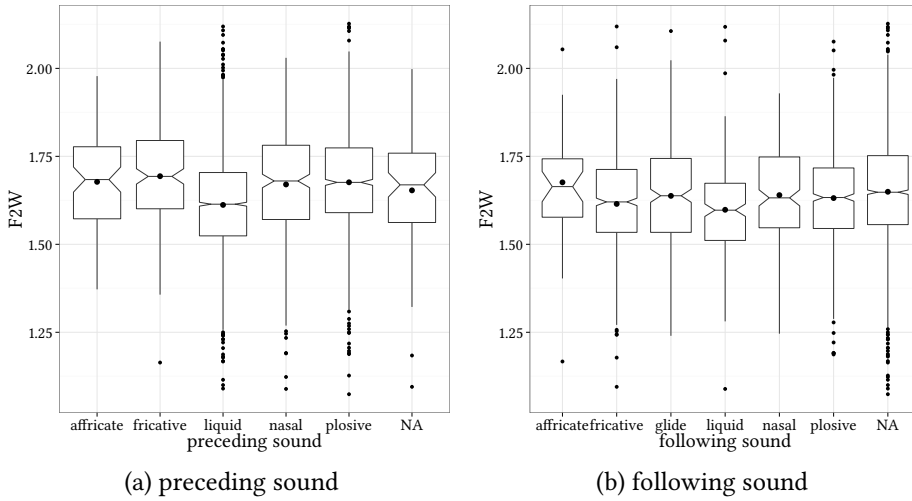


Figure 6.8: happy (F2) by preceding and following sound

names, 81 observations in total). Judging from this graph, it seems as if happy measurements following liquids were the odd ones out (with lower values of F2, on average), as the remaining means are much more similar to each other and the confidence intervals of the medians (illustrated by the notches) frequently overlap. Cases where happy is preceded by a liquid were also the ones (along with, to a lesser extent, preceding fricatives) that were found to be significantly different by the mixed linear effects regression model. It should be noted that this context (preceding liquid) includes regularly formed adverbs, and therefore, by itself, accounts for the majority of happy observations (2790 out of 4565, or 61.12%), which means that the statistical basis for this phonological environment is considerably larger than for the others. Furthermore, high frequency words such as *very* or *really* are to be found in this category, so it is not unlikely that phonological context is here mixed up with other features such as duration (see below).

Interestingly, a *following* liquid (at the beginning of the next word) seems to have a very similar effect on F2 measures as a preceding one (lowering of F2 in this context has been attested before, cf. [Lehiste 1964: 26](#)). The corresponding box plot for F2 by following phoneme (Figure 6.8b) shows happy to be somewhat more central (lower F2) when a liquid follows. This is in line with the regression model (which found a negative correlation coefficient for this context). The significantly negative coefficient (in the model) for a following fricative is less obvious

in this figure: while mean and median are somewhat lower, they do not appear to be significantly so (cf., for instance, the partially overlapping confidence notches of “fricative” and the neighbouring “glide”). The same holds for happy observations that are followed by an affricate. The mean in this category is higher than for other following consonants (corresponds to the positive coefficient found in the regression), but there seems to be a lot of noise in the raw data, which results in quite a large confidence interval (see the notches of the “affricate” box).

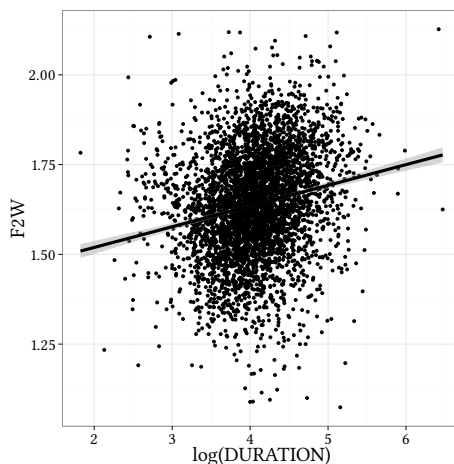


Figure 6.9: happy (F2) by duration

Let us now turn to vowel duration. The general effect is the same as for F1: longer vowel duration favours more peripheral vowel quality. In the case of F2, this translates into higher values (more advanced happy realisations). Figure 6.9 visualises the relationship in a regression plot, where log-transformed duration is marked on the x-axis and estimated F2 is found on the y-axis. While Figure 6.2 and Figure 6.9 cannot be directly compared (due to different scales), the mixed-effects models support the impression that the effect of duration on F2 is both stronger and more significant than on F1: the slope of the regression line is steeper for F2, the correlation coefficient is almost three times as big, and the p-value considerably smaller. With respect to what has been said about F2 values by preceding sound in particular, it seems worthwhile to quickly check whether vowel duration might have confounded the results reported in the preceding paragraphs.

Table 6.6 summarises mean durations of happy depending on which consonant precedes or follows. It is striking that happy realisations following liquids

Table 6.6: happy: durations by phonological environment

position	affricate	fricative	liquid	nasal	plosive	glide
preceding	79.84	70.35	62.44	62.98	70.45	NA
following	61.01	52.84	64.43	57.03	52.29	64.1

and nasals are considerably shorter, while a preceding affricate seems to have a lengthening effect. This might go some way towards explaining why happy was found to be significantly more centralised following a liquid: the effect might be due to shorter duration rather than the consonant that is found before the vowel. In this case, however, an explanation would be needed as to why a preceding nasal does not have a significant impact on F2 of happy, despite the fact that vowel duration is similar in these two contexts. Likewise, the higher duration of happy following an affricate should, but did not result in significantly fronter realisations (though this might be due to the small number of observations, cf. §6.1.1.2). When we look at the durations by following consonant, the picture becomes even messier. Before affricates happy is again slightly longer, but actually not as long as before liquids. A following affricate was found (in the regression) to be a factor favouring fronter happy realisations (in line with the duration values), but a following liquid actually has a *centralising* effect in the same model, even though happy is even longer in this environment. At least as far as the following sound is concerned, phonological environment thus does not seem to be confounded by vowel duration. At this point, no straightforward explanation presents itself as to why happy is significantly fronter in some contexts. Frequency of the keyword (as hinted at above) could be an option, but this factor was eliminated in the regression as non-significant. With respect to happy realisations in different phonological environments my account should therefore be considered more descriptive than explanatory.

6.1.2.3 Age and class

I will now turn once more towards the social predictors. Contrary to the F1 dimension, age of participant is not a significant main effect in the regression model estimating F2 of happy. The raw data, on the other hand, paint a different picture. The box plot (Figure 6.10) shows an increase in mean F2 from the oldest to the middle-aged group: happy has thus become fronter in this time frame. While the difference between groups is comparatively small, it is nonetheless highly

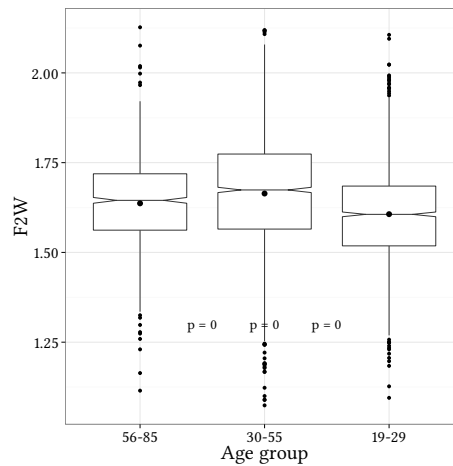


Figure 6.10: happy (F2) by age

significant ($t(2321.047) = 5.117$, $p < 0.001$). From the middle to the young group, however, the vowel does not become even more front (i.e. “more Scouse” as has been hypothesised), but rather it is centralised again. The youngest speakers in my sample do not only have happy realisations that are significantly more central than those of their parents’ generation ($t(3574.064) = -11.594$, $p < 0.001$). They use, in fact, variants which are, on average, even more retracted than those of the oldest speakers, a difference which is, once again, highly significant ($t(2190.006) = -5.655$, $p < 0.001$). Just as for F1, I have to conclude that the younger the speaker, the less Scouse – in this case front – happy is. That being said, it should be kept in mind that age of participant is no longer a significant fixed effect on its own once other factors and random variation due to carrier word are considered. The same goes for class of subject, even though this factor only just about fails to cross the 5% threshold. The interaction of age and social class, however, is a significant fixed effect in the regression model. Figure 6.11a and Figure 6.11b illustrate the relationship.

Figure 6.11b shows separate regression lines for the three age groups, which are as usual coded by line type: solid for the oldest, dotted for the middle-aged, and dashed for the youngest speakers. Estimated F2 is marked on the y-, and social class on the x-axis. This graph shows two interesting things: (1) Class does not only have an effect that is different in degree in the three age groups, but one whose direction is completely reversed in the youngest speakers. For both the old and the middle group, working-class speakers use more advanced happy

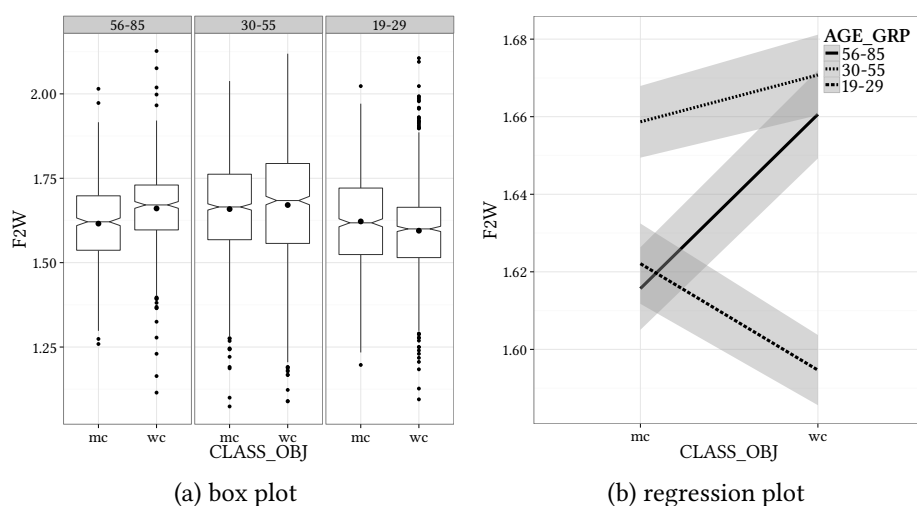


Figure 6.11: happy (F2) by age and class

variants. Young working-class speakers, on the other hand, have realisations that are more retracted than those of their middle-class counterparts. (2) Which age groups are different from each other depends on which social class one looks at. In the middle class, the oldest and the youngest speakers appear as one large group, without any significant differences between them ($t(1179.969) = 0.836$, $p = 0.403$), but one which does differ greatly from the middle group – compared to young ($t(1446.646) = -5.289$, $p < 0.001$), and old group ($t(1127.729) = 6.292$, $p < 0.001$) –, where realisations are a bit more front. Among the working-class speakers, on the other hand, the old and the middle-aged participants form a group, albeit one that shows somewhat more variation ($t(1192.156) = 1.219$, $p = 0.223$), which is now significantly different from the youngest speakers – compared to middle ($t(1659.468) = -10.25$, $p < 0.001$), and old group ($t(956.457) = -8.952$, $p < 0.001$) –, who use more retracted variants of happy.

The box plot next to this regression plot (6.11a) shows the same interaction, but the focus is now more on social class. There are three panels which illustrate the differences between middle- and working-class speakers within each age group separately. On the left-hand side one can see that in the oldest age group, working-class subjects have fronter realisations of happy than the middle-class speakers of the same age group. The difference looks highly significant (the confidence interval notches do not overlap), and indeed it is ($t(929.854) = -5.653$, $p < 0.001$). In the group of participants who are between 30 and 55 years

old, the mean of working-class subjects is also marginally higher, but this time there seems to be much more within-group variation and the two social classes are no longer significantly different from one another ($t(1668.574) = -1.663$, $p = 0.097$). When we look at the youngest speakers, we see again that the trend has reversed (working-class speakers in this subsample have more centralised variants of happy showing in lower F2 averages), variation – particularly among working-class subjects – has decreased and the difference between classes is once more statistically robust ($t(1443.946) = 3.878$, $p < 0.001$). There is thus no straightforward interpretation for either social class or age of participant, because their effects depend on each other not only in terms of degree, but also in direction. Notwithstanding the slightly confusing picture that these two factors create, one thing seems to be clear when we look at both the means in Figure 6.11a and the position of the dashed and dotted lines in Figure 6.11b: the youngest subjects in the sample have more retracted happy realisations than speakers of their parents' generation. They cannot be said to be more Scouse than their parents or grandparents.

6.1.2.4 Style and gender

The second significant interaction in the mixed-effects model, style X gender, is not really simpler in nature. The box plots in Figure 6.12 illustrate within-style comparisons by gender, but they also show style differences within genders. For female speakers (boxes on the left-hand side of each panel), the three styles 'read-

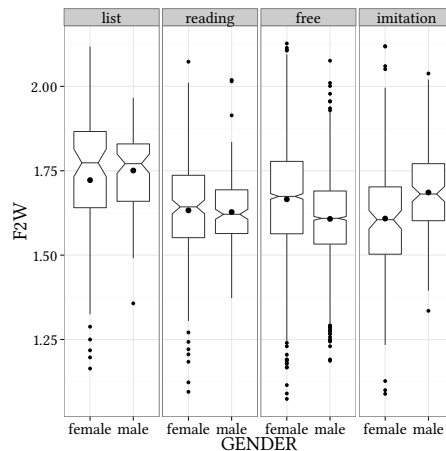


Figure 6.12: happy (F2) by style and gender

ing’, ‘free’, and ‘imitation’ have relatively similar realisations. This is especially true for reading and accent performance, where means are very close and notches overlap. Table 6.7 summarises t-tests on the raw data for all possible comparisons of style, separated by gender of subject. The p-value of 0.22 for reading-imitation (females) confirms that these two styles are not significantly different, whereas spontaneous speech differs from both reading ($p = 0.008$) and imitation ($p < 0.001$) in a statistically robust way. Realisations of happy elicited via the word list show slightly more variation around the mean, but are nevertheless clearly distinct from all three remaining styles as both the box plot and the t-tests (cf. Table 6.7) suggest.

Table 6.7: happy (F2): t-tests of style by gender

test	women			men		
	t	df	p	t	df	p
list-reading	3.396	113.708	< 0.001	8.343	165.401	< 0.001
list-free	2.370	79.086	0.020	11.426	94.870	< 0.001
list-imitation	3.994	141.852	< 0.001	3.984	197.933	< 0.001
reading-free	-2.661	244.712	0.008	2.362	234.779	0.019
reading-imitation	1.229	260.601	0.220	-4.279	279.877	< 0.001
free-imitation	3.482	146.933	< 0.001	-7.047	158.739	< 0.001

When we look at male speakers (right-hand boxes in Figure 6.12), accent performance and reading realisations are now clearly distinct ($p < 0.001$), and reading and spontaneous speech, in turn, look as if their means were comparable. A t-test on the raw data, however, does find this difference to be significant ($p = 0.019$), albeit slightly less so than those that hold between the other styles. The rightmost column of Table 6.7 indicates that (in the raw data!) men actually keep all four registers distinct. If we now consider gender differences within the individual registers, one thing that immediately strikes the eye is that, contrary to the other three, the difference in ‘reading’ seems negligible, which suggests that women and men do not have different happy realisations (with respect to F2 at least) in this speaking style.

The second panel from the left visualises gender differences for the reading passage only: the larger box for the female subjects indicates that there is more variation in this subgroup, but the means of men and women (marked by black dots) are indeed almost on the same level. Furthermore, the confidence intervals (notches) of the two boxes overlap, so this difference really seems to be non-

significant, a suspicion which is confirmed by a t-test ($t(353.108) = 0.338$, $p = 0.736$). Essentially the same holds true when subjects read out a word list (left-most panel): at first glance, men seem to use fronter happy variants than women (their mean is higher than that of females). However, the medians (black horizontal bars) of both subgroups are virtually identical and the notch of the ‘male’ box falls well within that of its ‘female’ counterpart, both of which suggest non-significance. A t-test again corroborates this interpretation ($t(113.484) = -1.07$, $p = 0.287$).

In spontaneous speech (‘free’ panel), on the other hand, there is a very clear and pronounced difference between female and male speakers. Women have a considerably higher mean F2 than men, and while they also exhibit quite a bit of variation (as illustrated by the vertical extent of the box), mean and median are almost identical for both the female and the male subjects, and the confidence intervals are very small and clearly distinct. The relevant t-test provides further evidence that women really do use more advanced happy variants than men in spontaneous speech ($t(3169.606) = 12.219$, $p < 0.001$). When participants are asked to imitate or perform a strong Scouse accent, this trend reverses completely, as is obvious from the rightmost boxplot in Figure 6.12. In this register, women’s mean F2 is lower, i.e. they produce more *retracted* realisations of happy than men, a difference which is highly significant ($t(232.688) = -4.024$, $p < 0.001$). Gender is thus only a significant factor in spontaneous speech and accent performance, but not for reading of a text or a word list. When it does play a role, its effect is not uniform: women are ‘more Scouse’ in free speech, but ‘less Scouse’ in the accent imitation task.

6.1.2.5 Style shifting

Before concluding this section, we will have a closer look at style and its (lacking) relation to age. I am aware of the fact that the linear mixed-effects regression did not find a significant interaction of these two predictors. Just as for the F1 dimension of happy I will print and describe the relevant plot all the same, because, (a) in this particular case, it visualises the impact of style almost as well as a plot that does not include age at all, and (b) both plots will later serve as points of reference for variables where there *is* a significant interaction. Figure 6.13 is identical in design to Figure 6.7: style is marked on the x-, and F2 on the y-axis. Age group is coded by line type, and the whiskers visualise standard deviations.

For all three age groups, there is a sharp drop from the word list to reading and the other two styles (with the exception of ‘imitation’ in the old group, which looks as if it might not be significantly different from the word list values in the

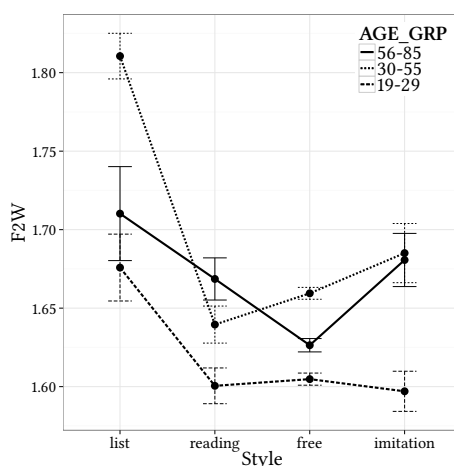


Figure 6.13: happy (F2) by age and style

same age group). This echoes what has been found for F1 of happy, and could, at least in parts, also be explained the same way: vowel duration is considerably higher when participants read out a word list (and also very similar in the other three registers; cf. Table 6.4), so the high values in this style are probably less to do with style shifting than with mere phonetic factors (although this does not explain why the difference seems to be particularly extreme for the middle age group).

One might be tempted to see some regularity in this graph (comparable to the u-shape in Figure 6.7), but the evidence is inconclusive. For the oldest speakers (where free speech marks the low point, and happy is more advanced both during reading and accent performance, which do not seem to differ significantly), a two-norms-approach might actually work pretty well: tense /i/ seems to be the target both in particularly local and particularly standard, or careful, speech. In the middle group, however, the three registers ‘reading’, ‘free’, and ‘imitation’ do not only apparently all differ significantly from one another (the standard deviations do not overlap), but happy is also more advanced in spontaneous speech than it is while reading a text – a fact, which does not go together very well with the assumption that a more formal register should pull realisations towards the tense standard, relative to a more relaxed (and therefore more central) starting point in spontaneous speech. For the youngest speakers in the sample, finally, this interpretation is completely out of the question. In this group, happy realisations are virtually identical with respect to F2 in reading, free speech, and accent

performance (standard deviations for these styles completely overlap). As far as the front-back dimension of happy is concerned, the youngest speakers that were interviewed for this study are not only, once again, the least Scouse (the dashed line is below the other two in all styles), but they also show the least awareness (*if* there is any to start with) of this feature as measured by style shifting.

6.1.3 Synthesis and Pillai scores (happy)

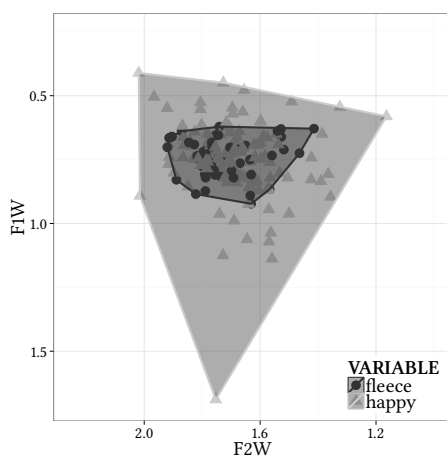
6.1.3.1 Overview

Moving away from awareness for a moment but keeping the focus on age, we will now zoom out a bit further and consider F1 and F2 measurements of happy together. The test variable happy will also be contrasted with FLEECE. This might provide a first hint as to whether just happy is changing in the younger generation, or whether other high front vowels are moving in the same (or another) direction as well. It should be noted, though, that FLEECE vowels were only measured in the reading and word list sections of the interviews (cf. §5.4.2). For the graphs reported on the following pages, the dataset of happy observations was therefore also reduced to the ones derived from reading and word list realisations to make sure like is compared with like.

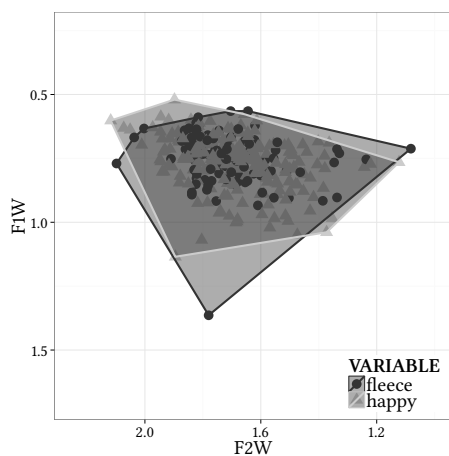
Figure 6.14 shows the spread of happy and FLEECE realisations for all three age groups separately. The three sub-plots are essentially ‘traditional’ F1-F2 vowel plots, which means that both the x- (F2) and the y-axis (F1) are inverted, so that fronter realisations are found to the left, and closer vowels towards the top of the graph. In all three plots, FLEECE realisations are marked by dark circles, and happy observations by light triangles. The dark and light polygons² connect the most peripheral pronunciations of each vowel and therefore define the total realisational space that the respective phonemes occur in with respect to the sample of this study.

Looking at the oldest speakers (Figure 6.14a), the first thing that probably strikes the reader is the outlier at around (1.75, 1.70), which results in a pronounced distortion of the happy polygon towards the bottom half of the graph. When this single (and rather extreme) realisation is ignored the light polygon becomes much more similar to those found in the middle-aged and the young group, with a lower limit at around F1 values of 1.15. The other interesting aspect is that FLEECE realisations are not only all to be found within the area defined by

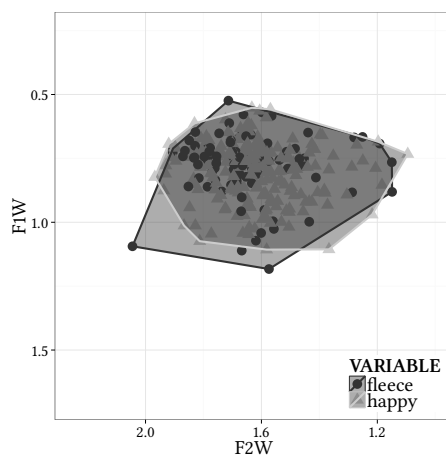
²The boundaries of the polygons were extracted from the dataset with the help of the R package ‘plyr’ (Wickham 2011).



(a) old speakers



(b) middle-aged speakers



(c) young speakers

Figure 6.14: happy-FLEECE: vowel space by age

happy, but that the realisational space of FLEECE (dark polygon) is also considerably smaller than that of happy. This indicates that FLEECE realisations are much more homogeneous among older speakers than happy pronunciations, although it has to be said it is possible that this result is at least in part due to the fact that there are comparatively few FLEECE observations in this age group to start with ($n = 52$). Since FLEECE is – visually – entirely included within the happy distribution, it is no surprise that a MANOVA returns a Pillai score near 0 and a p-value which clearly shows that the two vowels are (almost) completely merged (Pillai = 0.024, $F = 2.071$, $p = 0.129$).

In the middle-aged group (Figure 6.14b) there is a similar, if somewhat less extreme, outlier as in the sample of the old speakers: this time the dark FLEECE area seems distorted towards higher F1 values and begs the question whether this particular vowel should be excluded from the analysis. In any case, however, it is clear that the (horizontal) front-back extent of the FLEECE space is considerably larger than it was in Figure 6.14a, with or without the potential outlier. This is equally true for the light happy space, although the difference is less pronounced than for FLEECE. Nevertheless, it is obvious that there are happy realisations in the middle-aged group which are either slightly more front (towards the left of the graph) or more back (towards the right of the figure) than those found for the older speakers; the phonetic range of happy variants is thus a bit larger in the middle group. While the spread of FLEECE is much larger in the middle group, the overlap of FLEECE and happy seems to be just as complete as among the oldest speakers. A MANOVA confirms this impression by yielding very similar values as for the old speakers, which confirm that the vowel distributions are not significantly different from one another (Pillai = 0.003, $F = 0.53$, $p = 0.589$).

When we turn to Figure 6.14c, which represents the data collected from the youngest speakers, the picture changes only slightly. However, both FLEECE and, particularly, happy seem to be somewhat more retracted than in the middle-aged or the old speakers. This is evidenced by the lack of variants that are simultaneously very front and very high – like the ones found towards the upper left corner in Figure 6.14b and Figure 6.14a for the middle-aged and the old speakers, respectively. Just as for the middle group, however, the distributions of happy and FLEECE occupy essentially the same space and are not found to be significantly different by a Pillai test (Pillai = 0.015, $F = 2.298$, $p = 0.102$). The p-value of the MANOVA is close to a statistical trend, but this should not be overinterpreted: it is true that the difference between the two vowels is much closer to significance in the young group than in the middle-aged one, but the Pillai score is still virtually 0 in both cases, which means that even if there was a statistically ro-

bust difference among younger speakers, the degree of overlap between vowel distributions would be extremely high.

6.1.3.2 Age means

It *does* look as if the majority of happy realisations was marginally more retracted than most of the FLEECE variants in the young speakers, but since the centres of gravity of the vowel clouds are difficult to establish from Figure 6.14, it makes sense to consider a plot of the mean values (Figure 6.15). This graph ignores the spread and overall distribution and only indicates where the means, i.e. the centres, of the exemplar clouds are to be found. Two pieces of information can be extracted from this figure. Firstly, age groups cluster together: both mean FLEECE and happy of the old speakers are higher than either vowel is on average in the middle-aged group. The means for the latter group are, in turn, both more front and higher than either mean FLEECE or happy of the youngest subjects in the sample. The second point of interest in this graph is that the difference between the means of FLEECE and happy is smallest for the middle-aged realisations, larger for the observations pertaining to the older subjects, and most pronounced in the sub-sample of Liverpudlians aged between 19 and 29. This correlates with the fact that the p-value of the relevant MANOVA was largest for the middle group, smaller for the old, and close to the 0.10 threshold for the youngest

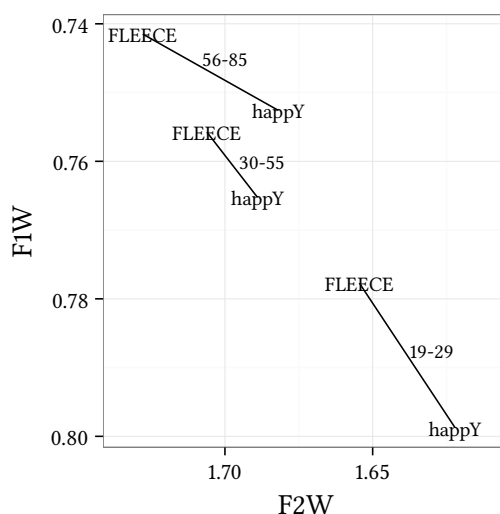


Figure 6.15: happyY-FLEECE: mean vowel position by age

speakers, which indicates that mean realisations of happy and FLEECE are indeed more robustly different among the old and the young speakers.

However, two caveats should be borne in mind: (1) Figure 6.14 and Figure 6.15 do not use the same scale – the means of FLEECE and happy are not identical, but the differences visible in Figure 6.15 are actually very small. (2) The previous point is corroborated by the MANOVA results, which indicate almost perfect overlap between overall distributions (Pillai scores near 0) and no significant difference between them (all p-values > 0.5). The available statistical evidence thus clearly supports the claim that FLEECE and happy are completely merged in all three age groups investigated. Of course, this does not (directly) touch on the primary issue of change in happy across the three generations. Figure 6.15 corroborates the findings reported for F1 and F2 separately (and on the basis of mixed linear effects regressions): happy becomes lower from the old to the middle-aged, and lower and *more* central from the middle to the young speakers.

All of this only refers to rather formal realisations of these vowels, since the results in this section are exclusively based on the reading passage and the word list. It remains to be seen whether the same conclusions would hold in spontaneous speech.

6.2 NURSE

6.2.1 F1 (NURSE)

6.2.1.1 Overview

Just as with happy results, the maximal model for NURSE F1 measurements exhibited severe collinearity ($\kappa = 38.1$). Separate regression models showed that both place and manner of articulation of the following sound ($\kappa = 34.25$), and place of following consonant and frequency of the carrier word ($\kappa = 21.22$) showed troubling or at least above average degrees of collinearity. Only one of these three, manner of articulation, was therefore retained. In a second maximal model, which neither included place of articulation of preceding and following sound, nor frequency of the keyword, collinearity was acceptable ($\kappa = 14.39$). The minimal adequate model (R^2 -equivalent = 0.314) that was then derived is shown below (Table 6.8).

Table 6.8: NURSE (F1): mixed linear effects regression

Fixed effects:	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	1.09	0.01	127.19	125.58	< 0.001	***
STYLElist	0.03	0.01	1078.12	3.61	< 0.001	***
STYLEread	0.00	0.01	1344.45	0.70	0.48	
STYLEfree	0.00	0.01	362.94	0.48	0.63	
AGE56-85	0.02	0.01	1475.42	3.19	< 0.01	**
AGE30-55	-0.01	0.01	1484.51	-2.86	< 0.01	**
GENDERf	-0.03	0.00	1477.09	-8.60	< 0.001	***
CLASSmc	0.05	0.00	1479.04	12.16	< 0.001	***
PREMANNERaffr	-0.05	0.02	74.59	-2.90	< 0.01	**
PREMANNERfric	-0.01	0.01	98.39	-0.82	0.42	
PREMANNERgli	0.01	0.01	65.20	0.54	0.59	
PREMANNERliq	0.02	0.02	118.12	0.82	0.41	
PREMANNERnas	0.02	0.02	377.34	0.95	0.34	
STYLElist:AGE56-85	0.02	0.01	1464.53	2.18	0.03	*
STYLEread:AGE56-85	-0.01	0.01	1465.50	-0.66	0.51	
STYLEfree:AGE56-85	0.01	0.01	1514.74	1.93	0.05	.
STYLElist:AGE30-55	-0.01	0.01	1468.34	-1.42	0.16	
STYLEread:AGE30-55	0.01	0.01	1470.78	0.80	0.42	
STYLEfree:AGE30-55	-0.01	0.01	1513.20	-0.89	0.37	
AGE56-85:GENDERf	-0.03	0.01	1475.15	-5.98	< 0.001	***
AGE30-55:GENDERf	-0.01	0.01	1475.93	-1.86	0.06	.
AGE56-85:CLASSmc	0.02	0.01	1471.18	3.15	< 0.01	**
AGE30-55:CLASSmc	-0.01	0.01	1475.87	-1.28	0.20	
GENDERf:CLASSmc	-0.01	0.00	1525.94	-2.80	0.01	**
STYLElist:AGE56-85:GENDERf	0.00	0.01	1464.29	0.25	0.80	
STYLEread:AGE56-85:GENDERf	0.02	0.01	1464.64	1.26	0.21	
STYLEfree:AGE56-85:GENDERf	0.02	0.01	1507.65	1.76	0.08	.
STYLElist:AGE30-55:GENDERf	0.00	0.01	1466.69	0.40	0.69	
STYLEread:AGE30-55:GENDERf	-0.01	0.01	1473.07	-1.07	0.28	
STYLEfree:AGE30-55:GENDERf	0.01	0.01	1506.49	1.63	0.10	
STYLElist:AGE19-29:GENDERf	0.01	0.01	1466.78	0.92	0.36	
STYLEread:AGE19-29:GENDERf	-0.00	0.01	1467.05	-0.03	0.97	
STYLEfree:AGE19-29:GENDERf	-0.02	0.01	1519.51	-2.30	0.02	*

6 Vowel production

STYLElist:AGE56-85:CLASSmc	0.01	0.01	1463.33	0.52	0.60	
STYLEread:AGE56-85:CLASSmc	-0.02	0.01	1464.04	-1.27	0.20	
STYLEfree:AGE56-85:CLASSmc	-0.02	0.01	1506.47	-1.85	0.06	.
STYLElist:AGE30-55:CLASSmc	0.02	0.01	1466.82	2.26	0.02	*
STYLEread:AGE30-55:CLASSmc	0.01	0.01	1469.86	1.27	0.20	
STYLEfree:AGE30-55:CLASSmc	-0.02	0.01	1497.92	-2.26	0.02	*
STYLElist:AGE19-29:CLASSmc	-0.00	0.01	1468.20	-0.42	0.67	
STYLEread:AGE19-29:CLASSmc	0.00	0.01	1468.92	0.19	0.85	
STYLEfree:AGE19-29:CLASSmc	0.02	0.01	1516.74	2.46	0.01	*
<hr/>						
Random effects:	(number of obs: 1568, groups: WORD, 137)					
Groups	Name	Variance	Std.Dev.			
WORD	(Intercept)	0.001	0.029			
Residual		0.011	0.103			

It is immediately obvious that this model contains several additional significant predictors when compared to the corresponding model of happy. Style, age group, gender, and social class are again all significant main effects. In addition, the model found significant interactions of style and age, age and class, age and gender, and gender and class. Towards the bottom of the model we see that the interaction that is of greatest interest for this study, style X age, further entered into significant three-way interactions with both gender and social class.

Non-social factors, on the other hand, seem to be less important than for happy, at least in relation to the social predictors: vowel duration, manner of following consonant, and frequency of the carrier word are all deleted as insignificant during model reduction. The only phonological predictor that is retained is manner of articulation of the preceding consonant, and even with this one there is only one level that is significantly different: NURSE realisations are higher (i.e. more standard) when they are preceded by an affricate. However, this only concerns a small minority of observations (97 out of 1770, or 5.48%), which are relatively equally distributed across all styles, so this predictor will not be analysed any further here.

6.2.1.2 Gender and class

Instead, we will move on to the interaction of gender and class. Since the interaction is only one of degree (see below) the relevant graphs do not add much to the discussion and are therefore omitted. In the group of middle-class speakers, men have a higher mean and median than women, which means that their NURSE realisations are lower. This difference is highly significant in the raw data ($t(958.928) = -8.555$, $p < 0.001$). For working-class subjects, we have essentially the same situation: women have a lower average and median F1, indicating that they use more central variants of NURSE ($t(751.346) = -4.357$, $p < 0.001$). Technically speaking, the difference between women and men is somewhat less robust in the working class group (cf. the t-values of the two tests), but for all practical purposes the relationship is the same in both social classes as both t-tests yield p-values that are below 0.001. The degree of the effect, however, is stronger in the middle class, meaning that there is a larger distance between the means of women and men in this class.

Conversely, the effect of social class is the same for both genders, although the difference between classes is smaller for women ($t(813.931) = 7.824$, $p < 0.001$), and more pronounced in the male group ($t(941.869) = 13.469$, $p < 0.001$). Once again, however, this distinction is rather fine-grained as class differences are highly significant for both genders. What is interesting, though, is that middle-class speakers actually use lower variants of NURSE than their working-class counterparts. This, in turn, holds for both genders, even though it is ever so slightly more pronounced in the male sub-sample.

6.2.1.3 Gender and age

The interaction of gender and age is somewhat more interesting in this respect. Looking at gender differences in the oldest speakers (left panel of Figure 6.16a) we find that women have considerably lower F1 values than men. The gender effect is highly significant ($t(395.594) = -5.823$, $p < 0.001$): men realise NURSE as a lower vowel than women. The same relation holds in the middle age group. Men again have higher F1 values on average than women. Judging from the box plot, which suggests less variation (smaller boxes) and smaller median confidence intervals (width of the notches), this difference is even more significant than for the oldest speakers. A t-test supports this impression ($t(805.568) = -11.767$, $p < 0.001$), although it has to be said that (a) this could simply be due to the fact that there are less data for the oldest speakers, and (b) the difference is already highly significant in the old group. The youngest subjects in the sample differ

6 Vowel production

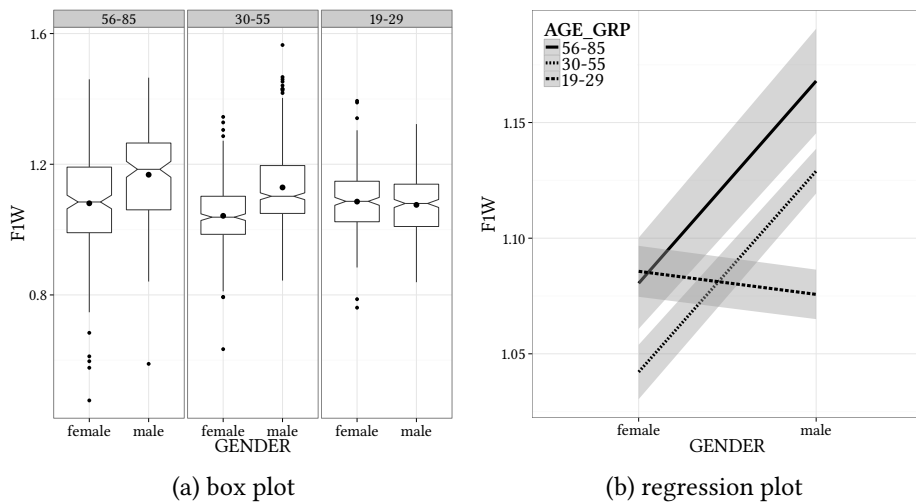


Figure 6.16: NURSE (F1) by gender and age

markedly from speakers of their parents' or grandparents' generation, because in this group there is no significant difference between male and female NURSE realisations ($t(537.587) = 1.277$, $p = 0.202$). Both means (and medians) are on a level which is almost perfectly intermediate between the means of middle-aged female and male speakers.

This non-significant difference between genders is also visible in Figure 6.16b, which has a (dashed) regression line for the youngest interviewees which is almost parallel to the x-axis of the plot. The comparatively steep positive slopes of the other two lines that stand for the old (solid) and middle-aged subjects (dotted), in turn visualise the gender difference that is already evident in Figure 6.16a. The regression plot also shows that all age groups seem to be significantly different from one another as far as the male speakers are concerned: the lines are clearly distinct and standard deviations (grey areas) do not overlap. One can reach the same conclusion based on the t-tests summarised in Table 6.9, which confirm that all three age groups are (highly) significant when the analysis is restricted to male subjects. In the female sub-sample, on the other hand, the middle age group is different from the other two, but the oldest and the youngest speakers do not differ with respect to the height of NURSE (cf. the t-tests in Table 6.9 and the standard error margins in Figure 6.16b). It looks thus as if male Scousers have constantly raised NURSE over the time period investigated here (if only to a very small extent in absolute terms), while female speakers first slightly raised NURSE

from the oldest to the middle-aged speakers, only to return to the starting point again in the youngest group. Since this starting point is statistically identical to the one the youngest male speakers have arrived at, the gender difference is therefore gone in this age group.

Table 6.9: NURSE (F1): t-tests of age by gender

test	women			men		
	t	df	p	t	df	p
old-middle	-3.292	337.596	0.001	-3.242	267.935	0.001
middle-young	5.819	566.832	< 0.001	-6.908	697.042	< 0.001
young-old	0.436	352.420	0.663	-7.637	268.135	< 0.001

6.2.1.4 Age and class

When it comes to the interplay of age and social class we have again a case where one might wonder why this interaction was found to be significant in the mixed linear effects regression. There are, once more, only differences in degree so the box and regression plots are not printed here (cf. §6.2.1.2). As reported above, middle-class speakers have higher F1 values (which translate to more open, i.e. more Scouse NURSE variants) than working-class Liverpudlians. This holds across all age groups and the difference is highly significant in the old ($t(403.191) = 5.927$, $p < 0.001$), the middle ($t(798.733) = 10.566$, $p < 0.001$), and the young group ($t(534.588) = 11.88$, $p < 0.001$), so an interaction does not seem ‘necessary’.³

Closer inspection of the figures nevertheless reveals tiny differences. The t-tests reported in Table 6.10 provide evidence that the middle-aged and the young group are not significantly different, neither in the working, nor the middle-class sub-sample. NURSE productions of the oldest speakers, on the other hand, seem to be distinct from the other two groups, both in the middle and the working-class sub-sample. In the raw data, p-values are only a bit higher for the working-class observations (but still below the 5% threshold, so NURSE variants in the oldest group *are* significantly different from the rest), but this slight deviation seems to be enough for the interaction to surface as significant in the mixed-effects model.

³A box plot based on the values predicted by the regression model (instead of the actually observed ones) was also generated to visualise the interaction once the random effects have been accounted for. This graph, however, did not look markedly different and is therefore not reprinted here either.

Table 6.10: NURSE (F1): t-tests of age by social class

test	middle class			working class		
	t	df	p	t	df	p
old-middle	-2.950	342.020	0.003	-2.064	266.401	0.040
middle-young	0.479	700.025	0.632	-0.119	610.856	0.905
young-old	-2.727	307.814	0.007	-2.133	267.448	0.034

Let us take a step back and briefly consider age of participant as a main effect before we investigate the interaction with speaking style. A box plot (Figure 6.17) clearly shows that the oldest subjects have a higher F1 for NURSE than both the middle and the young group. The difference between the old and the middle-aged group looks significant, and is indeed found to be so by a t-test on the raw data ($t(634.126) = -2.735$, $p = 0.006$). Equally significant ($t(1322.803) = -2.387$, $p = 0.017$) is the drop from the middle to the young group, even though the virtually identical medians and the notches of the boxes might suggest otherwise. Young speakers in my sample thus have a significantly more close NURSE realisation than speakers of the other age ranges.

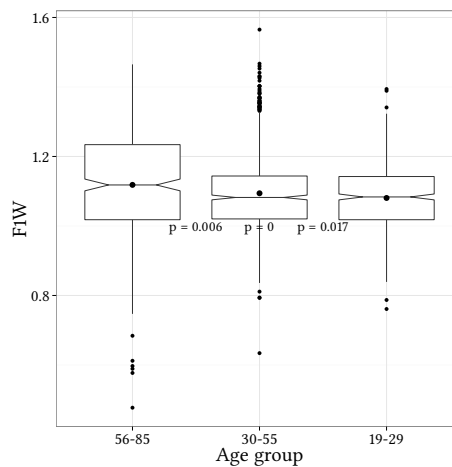


Figure 6.17: NURSE (F1) by age

6.2.1.5 Style shifting

The interaction of style and age (along with the two three-way interactions of style, age, and gender and class, respectively) are visualised by a number of line plots which are all structured similarly and were already used in §6.1.1 and §6.1.2: style is marked on the x-axis, F1 on the y-axis, and age group of participant is coded by line type.

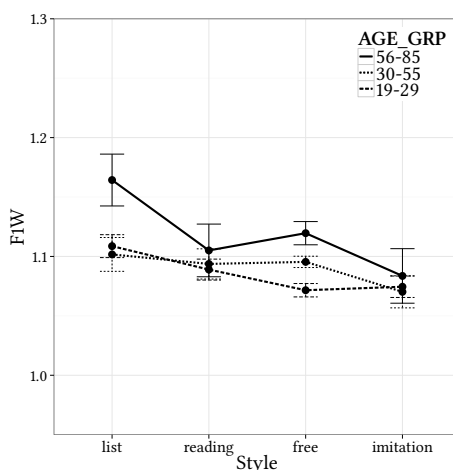


Figure 6.18: NURSE (F1) by style and age

Figure 6.18, which is based on the complete data set of NURSE observations, shows that differences between age groups are not really drastic. The young and the middle group have virtually identical values in three out of four speaking styles, only the oldest speakers are slightly more distinct. However, even that is mostly true for the word list. While reading a text and during accent performance all three groups have comparable F1 measurements. Only in spontaneous speech (which accounts for the clear majority of observations and therefore explains the results visualised in Figure 6.17) do all three groups have significantly different F1 means. No group shows systematic and significant style variation (note the often overlapping standard error bars between styles). Especially the lines of the young and middle-aged speakers look pretty level. If anything, style shifting can be found in the oldest participants, where there seems to be a more systematic downward trend from the left of the graph to the right (although the mean of ‘reading’ is a little off in this respect). However, just as in the other groups, the line potentially has a ‘wrong’ negative slope. If we were looking at Labovian style shifting, we would expect an *upward* slope, i.e. NURSE realisations becoming *more* Scouse from the word list to imitation, instead of the opposite.

6 Vowel production

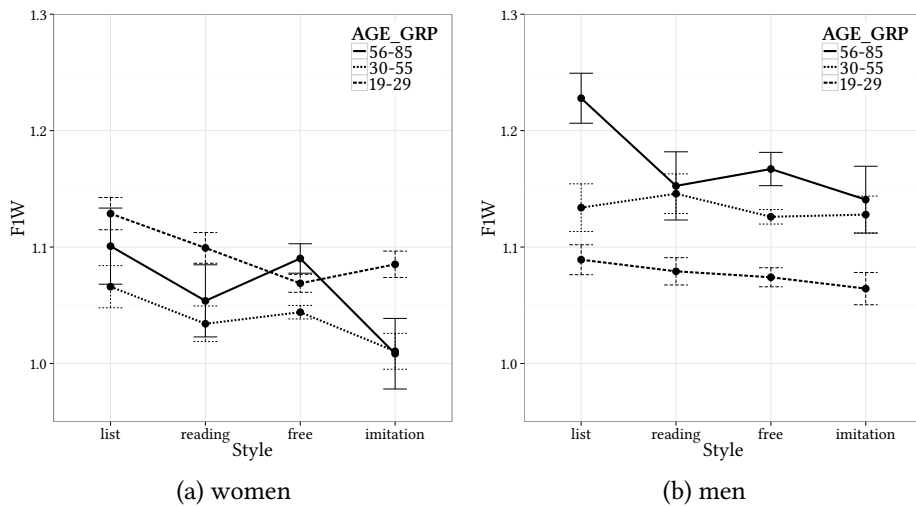


Figure 6.19: NURSE (F1) by style, age, and gender

Since the mixed linear effects regression found significant three-way interactions of style and age with both gender and social class, we should also look at the sub-samples defined by these two additional predictors. Figure 6.19a and Figure 6.19b show the same relationship of style and age as Figure 6.18, but they restrict the dataset to female or male subjects, respectively. It seems as if the slight downward trend that was visible for the oldest speakers only in Figure 6.18 can be found for females of *all* age groups (with the exception of the means in free speech, which are not really on the proposed line, but further up – for the middle and the old group – or down – for the young speakers – than they ‘should’ be). This cannot be called more than a subtle trend, however, as styles adjacent on the formality continuum are only rarely significantly different (‘reading’ and ‘free’ in the young, ‘free’ and ‘imitation’ in the other two groups). Also, the near-linear development is again in the wrong direction to qualify as style shifting, because realisations of NURSE become *less* Scouse in more informal contexts. If we look at the men in Figure 6.19b even this weak trend is gone. Middle-aged and young men have lines which are, for all practical purposes, flat, and the oldest speakers distinguish only the word list from the other three styles.

The interaction of style, age, and social class presents itself as somewhat more messy. Speakers in the old group can be said to be the ones where social class seems to play the least important role with respect to style shifting. In both the middle (Figure 6.20a) and the working class (Figure 6.20b), the general tendency

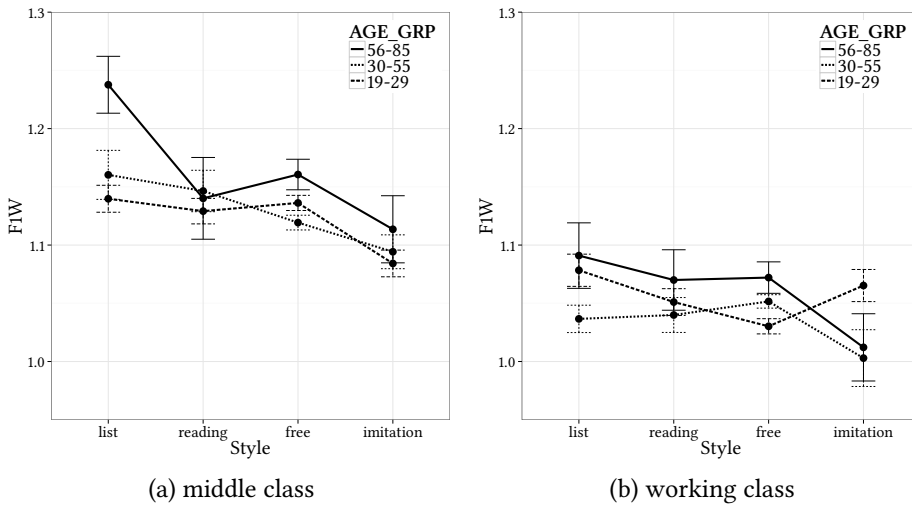


Figure 6.20: NURSE (F1) by style, age, and social class

for F1 values to decrease from most formal to least formal context is visible. Admittedly, vowels elicited by the word list are only (significantly) more open for old middle-class speakers, but, on the other hand, in *both* classes (a) realisations in the ‘reading’ and ‘free’ registers are not significantly different from each other, and (b) F1 of NURSE is (just about) significantly lower in accent performance (‘imitation’) than in spontaneous speech (‘free’).

Middle-aged subjects in the middle-class sub-sample generate the neatest version of the downward trend that should by now be familiar. First of all, F1 values for the dotted line decrease linearly, without exception, from left to right in Figure 6.20b. Second, if we lump together the word list and the reading passage (which do not seem to be different in a statistically robust way), this decrease is also significant from ‘reading’/‘word list’ to ‘free’, and from ‘free’ to ‘imitation’. Working-class speakers of the same age group do not show this regular downward trend. What is more, NURSE realisations in the registers word list, reading passage, and spontaneous speech are not – statistically speaking – different. When performing Scouse, variants are significantly higher, but only when compared to free speech; the overlapping error whiskers in Figure 6.20b indicate that ‘list’, ‘reading’, and ‘imitation’ are not significantly different for subjects aged between 30 and 55.

For the youngest group of speakers it is in the *middle* class where these three styles are not significantly different from one another. Accent performance, how-

ever, is then (highly) significantly different from all other registers. Young working-class speakers echo middle-aged middle-class participants in that they have a linear decrease in F1 from ‘list’ to ‘free’. In addition, both ‘list’ and ‘reading’ and ‘reading’ and ‘free’ are (at least marginally) significantly different. As a sort of reversal of the pattern found for middle-aged working-class speakers, there is then an increase towards the values found for ‘imitation’, a mean which is again only significantly different from spontaneous speech, but not the reading passage or the word list. There is thus a general similarity between Figure 6.19a and Figure 6.20a (downward trend more visible) and Figure 6.20b and Figure 6.19b (flatter lines), which is in line with an immense body of sociolinguistic research that has, time and again, found female speakers and those of higher socioeconomic classes to be more sensitive to sociolinguistically meaningful variables. If we zoom in, however, it becomes obvious that the middle-aged group exhibits a more linear pattern in the middle class (as expected), whereas the youngest speakers actually have a flatter line in this class than in the working class.

6.2.2 F2 (NURSE)

6.2.2.1 Overview

Compared to the model reported for F1 measurements of NURSE, the mixed linear effects regression of F2 obviously has a different dependent variable, but since there was an F2 measurement for every corresponding observation of F1, the independent variables (and their distribution within the dataset) are the same. Collinearity was therefore reduced in exactly the same way that was described at the beginning of §6.2.1. Model selection based on AIC scores and F-tests comparing nested models resulted in the minimal adequate model shown below (Table 6.11, R^2 -equivalent = 0.66).

Table 6.11: NURSE (F2): mixed linear effects regression

Fixed effects:	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	1.40	0.01	347.52	103.21	< 0.001	***
STYLElist	-0.01	0.01	1514.59	-1.33	0.18	
STYLEread	-0.03	0.01	1533.43	-5.22	< 0.001	***
STYLEfree	0.01	0.01	870.58	1.20	0.23	
AGE56-85	-0.03	0.01	1432.62	-5.14	< 0.001	***
AGE30-55	0.00	0.00	1447.88	0.38	0.70	
GENDERf	0.05	0.00	1516.12	18.82	< 0.001	***
CLASSmc	-0.06	0.00	1444.81	-18.59	< 0.001	***

DURATION	0.00	0.00	1506.54	3.16	< 0.01	**
PREMANNERaffr	0.07	0.02	71.47	3.44	< 0.001	***
PREMANNERfric	0.00	0.01	86.50	0.41	0.68	
PREMANNERgli	-0.07	0.01	62.72	-4.53	< 0.001	***
PREMANNERliq	-0.03	0.03	91.59	-1.07	0.29	
PREMANNERnas	0.01	0.02	159.94	0.57	0.57	
STYLElist:AGE56-85	-0.05	0.01	1425.23	-4.97	< 0.001	***
STYLEread:AGE56-85	-0.01	0.01	1419.97	-1.33	0.18	
STYLEfree:AGE56-85	-0.02	0.01	1493.50	-3.29	< 0.01	**
STYLElist:AGE30-55	0.02	0.01	1426.13	2.37	0.02	*
STYLEread:AGE30-55	0.02	0.01	1430.56	2.05	0.04	*
STYLEfree:AGE30-55	0.01	0.01	1493.35	2.63	0.01	**
STYLElist:CLASSmc	-0.02	0.01	1424.03	-3.49	< 0.001	***
STYLEread:CLASSmc	0.01	0.01	1425.34	0.99	0.32	
STYLEfree:CLASSmc	-0.01	0.00	1495.02	-2.31	0.02	*
AGE56-85:GENDERf	-0.02	0.00	1518.78	-5.49	< 0.001	***
AGE30-55:GENDERf	0.07	0.00	1512.01	22.03	< 0.001	***
AGE56-85:CLASSmc	-0.05	0.01	1431.94	-10.02	< 0.001	***
AGE30-55:CLASSmc	0.02	0.00	1449.85	4.97	< 0.001	***
GENDERf:CLASSmc	-0.02	0.00	1517.33	-8.76	< 0.001	***
STYLElist:AGE56-85:CLASSmc	0.04	0.01	1426.89	4.02	< 0.001	***
STYLEread:AGE56-85:CLASSmc	0.00	0.01	1419.93	0.52	0.61	
STYLEfree:AGE56-85:CLASSmc	-0.00	0.01	1481.98	-0.41	0.68	
STYLElist:AGE30-55:CLASSmc	-0.02	0.01	1427.41	-2.22	0.03	*
STYLEread:AGE30-55:CLASSmc	-0.01	0.01	1430.06	-1.78	0.08	.
STYLEfree:AGE30-55:CLASSmc	-0.01	0.01	1487.18	-1.03	0.30	
<hr/>						
Random effects:	(number of obs: 1568, groups: WORD, 137)					
Groups	Name	Variance	Std.Dev.			
WORD	(Intercept)	0.002	0.044			
Residual		0.008	0.091			

This model contains all the social factors as significant fixed effects which were already found to be significant predictors when modelling the F1 values of NURSE. The only exception to this is the three-way interaction of style, age, and gender which does not reach statistical significance in the regression of F2. Age, gender, and social class of participant, however, are all significant main effects. So is speaking style, which further acts as a predictor of F2 in two-way interactions with age and social class. Gender of speaker also interacts with social class and

age each. Age, finally, furthermore appears both in a significant two-way interaction with social class and a three-way interaction of style, age group, and social class. In addition to the social categories, phonetic and phonological predictors are more important than they were for the height dimension of NURSE: both the preceding consonant and the duration of the observed vowel have a significant impact on F2 values of NURSE. It is these two factors which we will look at first.

6.2.2.2 Phonological context

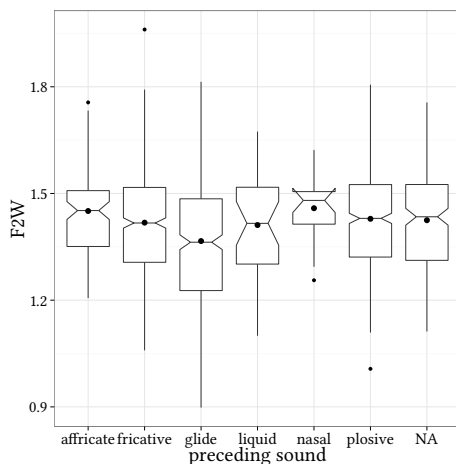


Figure 6.21: NURSE (F2) by preceding sound

The mixed linear effects regression found two levels of the factor (manner of) preceding consonant in particular to have a significant impact on F2 of NURSE. A positive correlation coefficient was calculated for cases where NURSE is preceded by an affricate (indicating that this phonological context favours fronter, more Scouse, realisations), whereas a negative coefficient (more central NURSE variants) was found for preceding glides. Realisations of NURSE that are preceded by a glide have, in fact, the lowest means in the raw data as Figure 6.21 shows. Measurements taken after an affricate have a mean that is higher than those of most other categories (with the exception of NURSE following a nasal). It is unclear why a preceding affricate should pull NURSE more to the front, but as has already been explained in §6.2.1, only a very small number of NURSE observations was made following an affricate anyway. This might mean that the result is somewhat shaky, but on the other hand it is striking that this phonological context pops up as significant again and again in this study.

Table 6.12: NURSE: durations by preceding consonant

affricate	fricative	liquid	nasal	plosive	glide
178.45	156.32	210.14	145.86	164.47	111.73

As far as F2 of NURSE is concerned, vowel durations, which are reported in Table 6.12, could give at least a hint about what might be going on, namely that the influence of the preceding sound is in fact confounded with duration. Some of the evidence is in conflict with this claim, but it is nonetheless striking that contexts where NURSE is preceded by an affricate have the second highest average duration in the sample, whereas NURSE is (by far) shortest when it follows a glide.

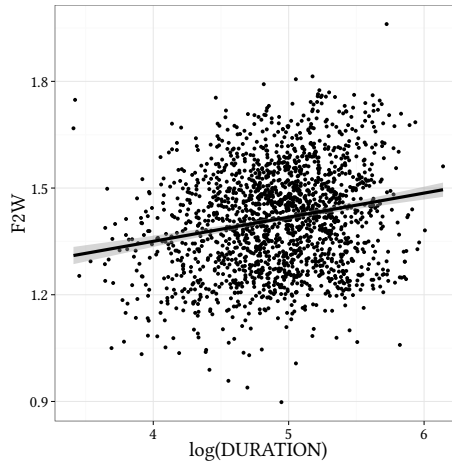


Figure 6.22: NURSE (F2) by duration

Figure 6.22 plots the log-transformed duration of NURSE on the x- and the F2 values on the y-axis. Just as with happy, longer vowel duration favours more peripheral realisations. It is therefore possible that the effect found for different preceding consonants is really due to duration. A further hint in that direction is provided by a regression of F2 on manner of preceding consonant and duration of the vowel only, which expressed medium collinearity ($\kappa = 10.48$), a fact that indicates these two predictors are not completely orthogonal and explain, if only to a small degree, the same part of the variation in F2. At least to a certain extent, the phonological effect of the preceding consonant in this sample can thus be seen as an artefact of a phonetic one, and, for this reason, will not be discussed here any further.

6.2.2.3 Gender and class

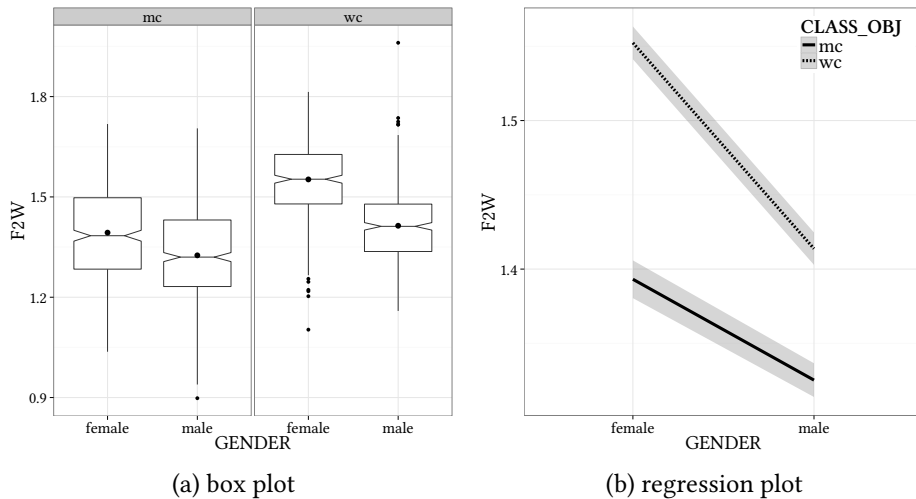


Figure 6.23: NURSE (F2) by gender and class

The first interaction we will look at is that of gender and social class of participant, which is illustrated in Figure 6.23a and Figure 6.23b. The two box plots show the difference between the two genders in the middle and the working class, respectively. Male speakers have a lower mean F2 than women in both classes. This is a surprising result because lower F2 values mean more central realisations and more central variants of NURSE are *less* Scouse variants of NURSE. In most sociolinguistic studies, however, men have been found to be *more* likely than women to use local variants of socially meaningful variables. Judging from the plot, the difference between genders is already highly significant in the middle-class sample (left panel) because not only are the means of women and men clearly distinct, but they are also virtually identical to the medians of the same category (which argues for normally distributed data), and the confidence intervals do not occupy the same space at all. A t-test confirms this interpretation ($t(901.122) = 7.779$, $p < 0.001$). It is also obvious, however, that the difference between women and men is much more pronounced in the working class (right panel): (a) The distance between the means is considerably greater, and (b) even the interquartile ranges (visualised through the size and position of the boxes) hardly overlap, let alone the confidence intervals ($t(791.427) = 17.513$, $p < 0.001$).

The fact that gender has a more drastic effect in the working class than in the middle class is also illustrated by the regression plot in Figure 6.23b, where

gender is to be found on the x-, and estimated F2 on the y-axis. A greater effect of gender should, in this graph, translate to a steeper slope of the regression line from ‘female’ to ‘male’, and this is precisely what we find when we compare the dotted (working class) to the solid (middle class) line. What we can also see is that middle-class speakers have lower F2 values than their working-class counterparts (the solid line is below the dotted one). This is not surprising because it means that middle-class speakers use less Scouse variants than working-class Liverpooldians, which is true for both female ($t(815.432) = -18.076$, $p < 0.001$) and male subjects ($t(933.018) = -11.305$, $p < 0.001$). Social classes are less distinct when we focus on male subjects only (the vertical distance between the regression lines is smaller), but it should be noted that this is a highly relative statement as the difference is statistically extremely robust for this sub-group of speakers, too. The significant interaction in the mixed-effects regression is thus not due to there being an effect in different directions in the sub-samples, but it is ‘merely’ an expression of the fact that the effect of gender (class) is stronger for working-class (female) subjects.

6.2.2.4 Gender and age

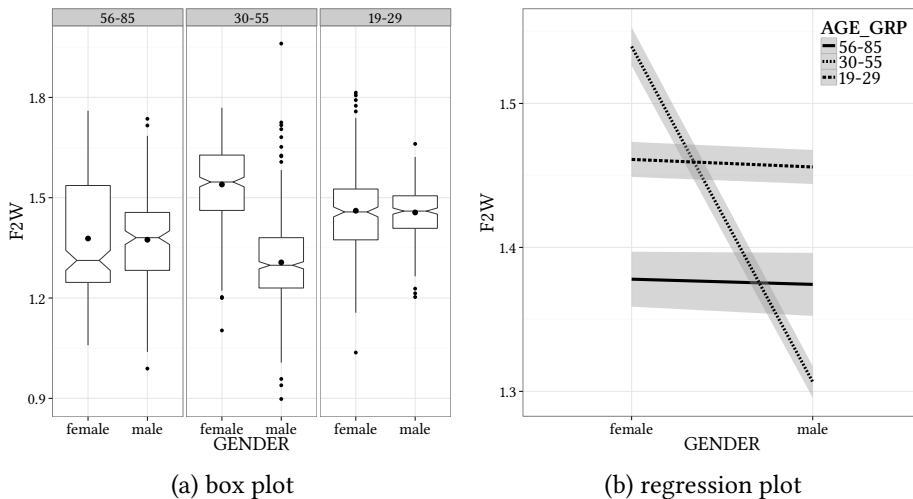


Figure 6.24: NURSE (F2) by gender and age

In the interaction with age, on the other hand, the effect of gender is clearly not the same in the different sub-samples. If we focus on the gender difference in the oldest subjects (left panel in Figure 6.24a), the medians (and the notches

that go with them) might suggest at first glance that male speakers have higher F2 than women, because the notches signal that the medians of the two groups are significantly different. The means, however, are virtually on the same level, because the female sub-sample is skewed towards the upper end of the scale (the arithmetic mean is considerably higher than the median due to comparatively extreme values between the second and third quartile). There is also quite a bit of variation in the female sub-sample (the box is almost twice as big as that of the old men). Indeed, a t-test reveals that NURSE productions of the oldest women and men in the sample are not significantly different from one another ($t(405.995) = 0.254$, $p = 0.8$). The same is true for the youngest speakers. Here, even the box plot does not suggest otherwise, because the data are more normally distributed for both genders (means and medians are very close to each other). The relevant t-test confirmed that, statistically speaking, NURSE is as front in the young women as it is in the young men ($t(440.579) = 0.6$, $p = 0.549$). In the middle-aged group, however, realisations are just as clearly distinct as they are identical for the other two age groups. Women aged between 30 and 55 use NURSE variants which are, on average, very noticeably fronter than those used by men of the same age group. The box plot shows that there is no overlap at all between the interquartile ranges, so the middle 50% of the data occupy completely separate areas on the scale for men and women. It is therefore not at all surprising that a t-test finds this difference to be highly significant ($t(710.778) = 26.419$, $p < 0.001$).

Table 6.13: NURSE (F2): t-tests of age by gender

test	women			men		
	t	df	p	t	df	p
old-middle	12.723	407.845	< 0.001	-6.203	308.955	< 0.001
middle-young	-7.794	567.234	< 0.001	20.528	760.911	< 0.001
young-old	6.378	419.726	< 0.001	7.804	260.802	< 0.001

In Figure 6.24b, the non-significance of gender in the oldest and the youngest speakers is expressed by the fact that both the solid and the dashed lines are essentially flat and parallel to the x-axis. The dotted line that stands for the middle-aged speakers, on the other hand, has a very pronounced negative slope, echoing that women in this group (surprisingly) use fronter NURSE variants than men. The other piece of information that can be extracted from this figure is that gender does not impact on which age groups are significantly different from one another. Both for female and male subjects regression lines are quite clearly sep-

arate and there is not even a hint of overlap for the grey standard error bands. Table 6.13 provides the data yielded by t-tests on the raw data, which found differences that were statistically robust for all combinations of age groups and within both genders, so this aspect is not influenced by gender. What *is* different for women and men is the positioning of age groups on the ‘Scouseness scale’: while for female speakers the middle group is the most Scouse, followed by the young and then by the old subjects, male middle-aged speakers have the most central NURSE variants, old speakers are in the middle, and the youngest men in the sample have the most advanced and therefore also the most Scouse realisations of NURSE.

6.2.2.5 Style and social class

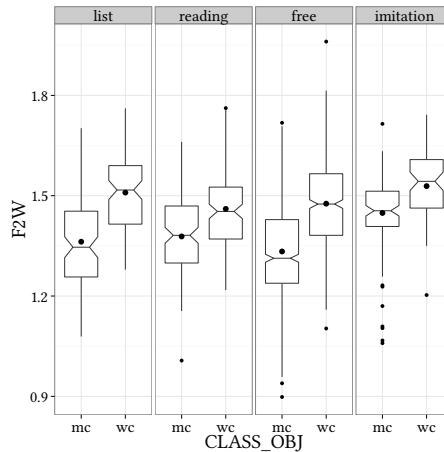


Figure 6.25: NURSE (F2) by style and social class

The interaction of style and social class is again one where we are looking at differences in degree more than in nature. If one examines the boxplots in Figure 6.25, which are sorted from most formal register (word list) on the left to least formal one (accent performance) on the right, it is clear that the pattern is pretty much the same in all of them. Boxes that visualise the data collected from middle-class interviewees (always the left one in each panel) are consistently lower than the corresponding ones for working-class speakers, which illustrates once again that working-class speakers have more Scouse NURSE variants than middle-class Liverpudlians (cf. Figure 6.23b). T-tests confirm for the means what the notches in the box plots do for the medians: the class difference is statistically significant for the word list ($t(188.089) = -8.096$, $p < 0.001$), the reading passage ($t(255.085) =$

-5.524, $p < 0.001$), spontaneous speech ($t(1087.269) = -17.359$, $p < 0.001$), and accent imitation ($t(153.537) = -4.744$, $p < 0.001$).

Table 6.14: NURSE (F2): t-tests of style by social class

test	middle class			working class		
	t	df	p	t	df	p
list-reading	-0.923	189.641	0.357	3.019	210.158	0.003
list-free	1.971	130.746	0.051	2.473	147.930	0.015
list-imitation	-4.894	192.492	< 0.001	-1.107	153.071	0.270
reading-free	3.870	209.460	< 0.001	-1.254	216.848	0.211
reading-imitation	-4.684	233.144	< 0.001	-4.039	157.841	< 0.001
free-imitation	-9.346	172.419	< 0.001	-3.684	101.284	< 0.001

Figure 6.25 also shows that while the class effect is generally the same for all styles, it is more pronounced in the reading list and free speech registers, than when people read out a text or put on a particularly strong Scouse accent. Both classes have the most Scouse realisations in the accent performance, but apart from that the order of styles is different (cf. Table 6.14 for tests of significance). For middle-class speakers ‘imitation’ is significantly more Scouse than ‘reading’, which is in second place and followed by the word list (which does not differ significantly from ‘reading’). The most central (least Scouse) variants are found in spontaneous speech, which is very close to forming a statistically robust contrast with the word list. When we look at working-class subjects, the order from most Scouse to least Scouse is: ‘imitation’, ‘list’, ‘free’, and ‘reading’. The first and the last two of these four are not significantly different from each other, so that statistically speaking we have a more Scouse block ‘imitation/list’, and a less Scouse pair ‘free/reading’. None of these orders corresponds to what would have been expected based on the traditional formality scale.

6.2.2.6 Age and social class

In addition to style, social class also interacts significantly with age group, but before we turn to this interaction we will first have a quick look at age as a main effect, where a very clear and simple picture emerges (Figure 6.26). F2 values increase significantly from the oldest to the middle-aged group ($t(914.871) = 2.558$, $p = 0.011$), and continue to rise from the middle to the youngest speakers ($t(1347.58) = 7.955$, $p < 0.001$). Middle-aged speakers have thus a fronter NURSE than the old-

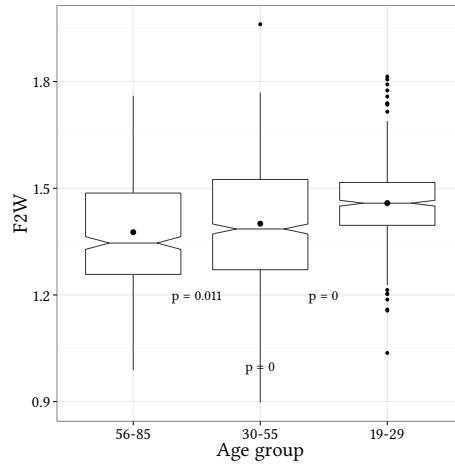


Figure 6.26: NURSE (F2) by age

est subjects in this sample, and speakers in the youngest group have, in turn, a significantly fronter vowel than Liverpooldians of their parents' generation. However, things are slightly more complicated, as an analysis of the age X social class interaction reveals.

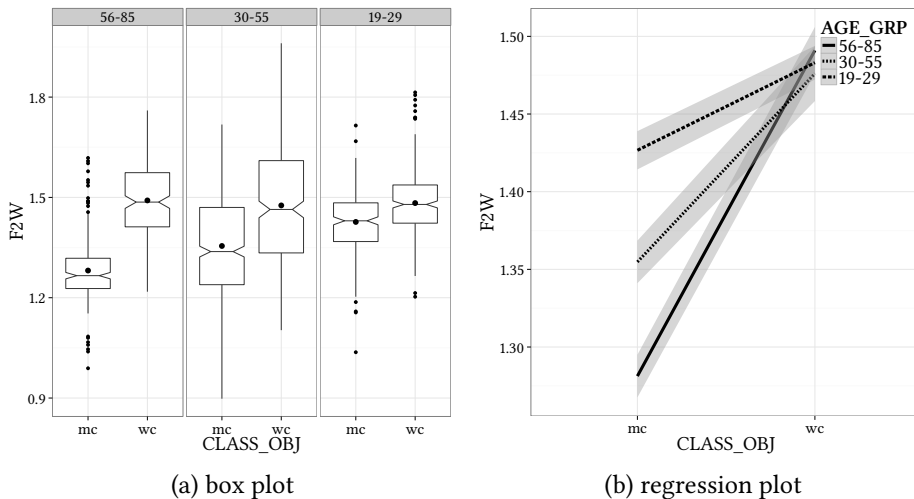


Figure 6.27: NURSE (F2) by age and class

When we look at class differences in the three age groups separately, we arrive at the picture represented by the box plots in Figure 6.27a. In all three groups, the means of middle-class speakers are lower than that of the working-class interviewees. For the oldest speakers (left panel) the difference is particularly pronounced. Both means and medians are different in a statistically robust way: the interquartile ranges of the two boxes do not overlap, and a t-test also finds this difference to be significant ($t(362.001) = -19.881$, $p < 0.001$). In the middle group there is a greater amount of variation (both boxes are larger) and both means and medians are closer to each other. As a consequence, the interquartile ranges do overlap, but middle-class and working-class speakers do all the same differ in a highly significant way in this age group as well ($t(615.171) = -10.466$, $p < 0.001$). When the dataset is restricted to the youngest speakers the difference between social classes is further reduced. At the same time values are more homogeneous in this age group – visually represented by the vertical extent of the boxes, which is smaller than in the middle group and comparable to the one found for the oldest speakers. While less pronounced, the difference between middle and working-class speakers is therefore just as significant for the youngest subjects as it is in the other groups ($t(521.157) = -6.815$, $p < 0.001$).

Table 6.15: NURSE (F2): t-tests of age by social class

test	middle class			working class		
	t	df	p	t	df	p
old-middle	7.842	645.309	< 0.001	-1.189	482.408	0.235
middle-young	7.920	698.035	< 0.001	0.638	507.159	0.524
young-old	16.507	452.539	< 0.001	-0.784	349.296	0.434

Figure 6.27a also shows that it is due to the middle-class speakers that the social class difference decreases linearly from the oldest to the youngest speakers. Working-class realisations are, in fact, remarkably stable over the three generations investigated here. For middle-class speakers, on the other hand, NURSE variants become consistently fronter from the oldest to the youngest speakers. In Figure 6.27b, the decreasing distance between social classes is visible in the slopes of the regression lines, which become flatter from the old (solid) to the middle-aged (dotted) and young speakers (dashed). This graph also provides support for the other point made above. For middle-class speakers (on the left-hand side of Figure 6.27b), all three age groups are different from each other (cf. the three relevant t-tests in Table 6.15, which all yielded significant results). In the

working-class sub-sample, all regression lines occupy essentially the same place, and all standard deviations overlap (dark grey area). T-tests on the raw data (cf. again Table 6.15) also confirm the claim voiced above that age differences only exist in the working, but not the middle class.

6.2.2.7 Style shifting

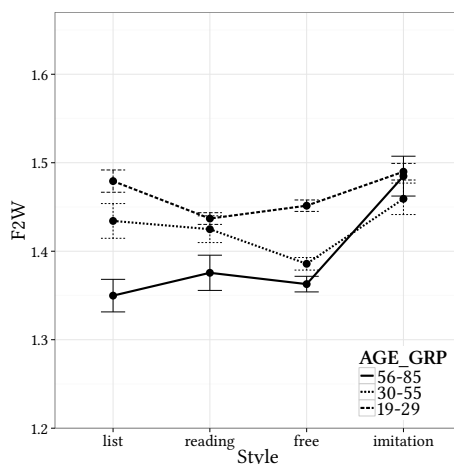


Figure 6.28: NURSE (F2) by style and age

The last combinations of predictors that can shed light on the (change in) salience of NURSE are the two-way interaction of style and age, and the three-way interaction of style, age, and social class. The relationship of style and age of participant is visualised in Figure 6.28. For the oldest group (solid line) there seems to be some movement from the word list to the reading passage to spontaneous speech, but in fact the standard error whiskers indicate that, statistically speaking, these three styles have identical NURSE realisations. Only the extreme rise when actively performing Scouse results in a mean that is significantly different from the other three registers. This sharp increase in F2 values from free speech to accent imitation is also present (and just as significant) in the middle-aged group. Spontaneous realisations in this age group are also significantly different from the ‘reading’ register, but, interestingly, NURSE is *fronter* – more Scouse – in this more formal register. The same goes for the word list, a register which is not significantly different from reading for these subjects. F2 values in the middle age group thus strongly hint at hypercorrection (cf. §6.2.3). Speakers in the young group, finally, exhibit a pattern which is closest to what we would expect if we

replaced
‘tell us
some-
thing
about’

6 Vowel production

assume there to be style shifting: F2 increases from reading a text to free speech and from spontaneous speech to accent imitation; both these increases are (just about in the former case) statistically robust. The only thing that does not fit into this frame is that realisations when reading out the word list are actually just as front as when these speakers imitate a particularly strong Scouse accent (no significant difference). In the most formal register young speakers' behaviour is therefore opposite to what would be expected.

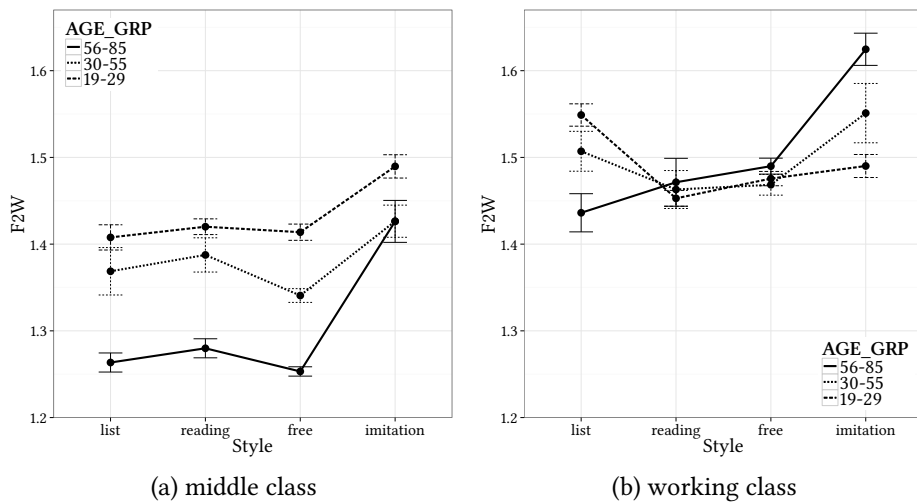


Figure 6.29: NURSE (F2) by style, age, and social class

Since the mixed linear effects regression also found a three-way interaction of style, age, and social class of participant, we will end this section with an analysis of two figures that illustrate the interaction of style and age for middle-class (Figure 6.29a) and working-class subjects (Figure 6.29b), respectively. Class seems to play the least important role for the oldest interviewees in the sample. Middle-class speakers in this age group have an almost flat line for the first three registers: there is a slight but significant drop from 'reading' to 'free', but word list realisations are neither significantly different from those in the reading passage, nor from the variants that are found in spontaneous speech. Accent imitation is significantly (and drastically) fronter than the other three registers. The solid line in Figure 6.29b suggests a more steady increase of F2, but 'list' is not significantly different from 'reading', which in turn is not significantly different from 'free' ('list' and 'free', however, are) – statistically speaking, this line can be thought of as flat for the first three styles, just as in Figure 6.29a.

The dotted line (middle-aged subjects) in Figure 6.29a is very similar to the one found for the entire dataset in Figure 6.28. Word list and reading passage are not significantly different, F2 drops significantly in free speech, and rises significantly again in accent performance. Working-class speakers in this age group, on the other hand, have statistically identical NURSE variants in free speech and while reading a text. NURSE is again fronter when people imitate a strong Scouse accent, but the vowel is actually just as advanced in the word list, the register which was supposed to elicit more standard pronunciations.

For the youngest speakers in the sample the class distinction is most interesting. Middle-class speakers exhibit a statistically completely level line for the first three styles, followed by the same rise in F2 for ‘imitation’ as in the old and the middle-aged subjects. Working-class speakers, on the other hand, do not significantly distinguish free speech from accent performance. They do, however, have significantly more retracted variants when reading out a text (as compared to spontaneous pronunciations in connected speech). Furthermore, when reading out a word list young working-class speakers actually have more advanced (more Scouse) NURSE realisations than in any other register. The shifting found in Figure 6.28 for the youngest speakers thus seems to be a true combination of these two diametrically opposed patterns: The significant fronting of NURSE during accent performance is due to middle-class speakers, whereas their working-class counterparts contribute the fronting in the most formal word list style (and also drive the rise in F2 from ‘reading’ to ‘free’).

6.2.3 Synthesis and Pillai scores (NURSE)

6.2.3.1 Overview

Up until now, the analysis has exclusively focused on the realisation of NURSE, despite the fact that we are dealing with a vowel *merger* here. Since (near-standard) SQUARE is generally considered to be the target of this merger in Liverpool, said focus on NURSE does not seem unjustified, since it is mostly in the latter one that variation is to be expected. Nevertheless, the other member of the pair should not totally be neglected. After all, it could well be that both NURSE and SQUARE are, for example, fronted by a particular group of speakers. This *might* still make them sound more Scouse (because NURSE would be moving further away from the standard), but not necessarily so because it is primarily the merger of these two vowels that is considered a characteristic feature of the accent. While constraints of space do not permit a detailed analysis of SQUARE (such as the one provided for NURSE in §6.2.1 and §6.2.2), the remainder of this chapter will at

least have a brief look at the realisational spaces of both *NURSE* and *SQUARE* in the different age groups, and it will also consider Pillai scores calculated on the basis of these distributions.

Table 6.16: *NURSE-SQUARE*: Pillai scores by age and style (groups)

style	old		middle		young	
	Pillai	p-value	Pillai	p-value	Pillai	p-value
word list	0.021	0.436	0.015	0.333	0.025	0.150
reading	0.026	0.315	0.017	0.208	0.034	0.051
free	0.073	< 0.001	0.050	< 0.001	0.025	0.007
imitation	0.016	0.533	0.007	0.727	0.014	0.398
total	0.046	< 0.001	0.031	< 0.001	0.013	0.004

Table 6.16 summarises the results of a number of MANOVAS, which were carried out in each age group and in each speaking style separately. The last line in this table (which reports the tests performed on all *NURSE* and *SQUARE* observations within the respective age group, pooled across style) tells us that *NURSE* and *SQUARE* distributions are (highly) significantly different in all three age groups investigated – all p-values are well below the 5% threshold. However, these values must be interpreted in connection with the Pillai scores that go with them, and which are just as homogeneous as the p-values: all of them are close to 0 (remember that Pillai scores can have values between 0 and 1), which indicates almost perfect overlap between the two (merged) vowel distributions. It is therefore more than likely that the low p-values are simply due to the large amount of data that are available when the observations are pooled across speaking styles, because in a sufficiently large dataset *any* difference will turn out to be statistically significant. We can therefore say that all three generations of speakers have completely merged *NURSE* and *SQUARE* distributions.

Judging from Table 6.16, style does not seem to make much of a difference either. It is true that some Pillai scores calculated for a particular register are almost twice (or, in the case of the youngest speakers, even thrice) as large as the total score in the last line, but even so all Pillai scores are (considerably) below 0.1, meaning they hardly differ from the total scores in absolute terms. This time, p-values support the idea that distributions are almost perfectly merged: With the exception of the ones pertaining to spontaneous speech (which accounts for the lion share of the data in each age group), no p-value is below the 5% threshold. All age groups seem to have merged vowel distributions in all four speaking styles for which data were collected.

6.2.3.2 Old speakers

It still does make sense to have a closer look at the individual distributions. Pillai scores allow statements about the degree of overlap between two distributions, but they say nothing about the exact shape, the location, or the centre of gravity of these distributions. Figure 6.30 represents the realisations spaces of *NURSE* and *SQUARE* in the different styles for the oldest speakers in the sample. Here and in Figure 6.31 and Figure 6.32 the vowel spaces that *NURSE* occurs in are marked

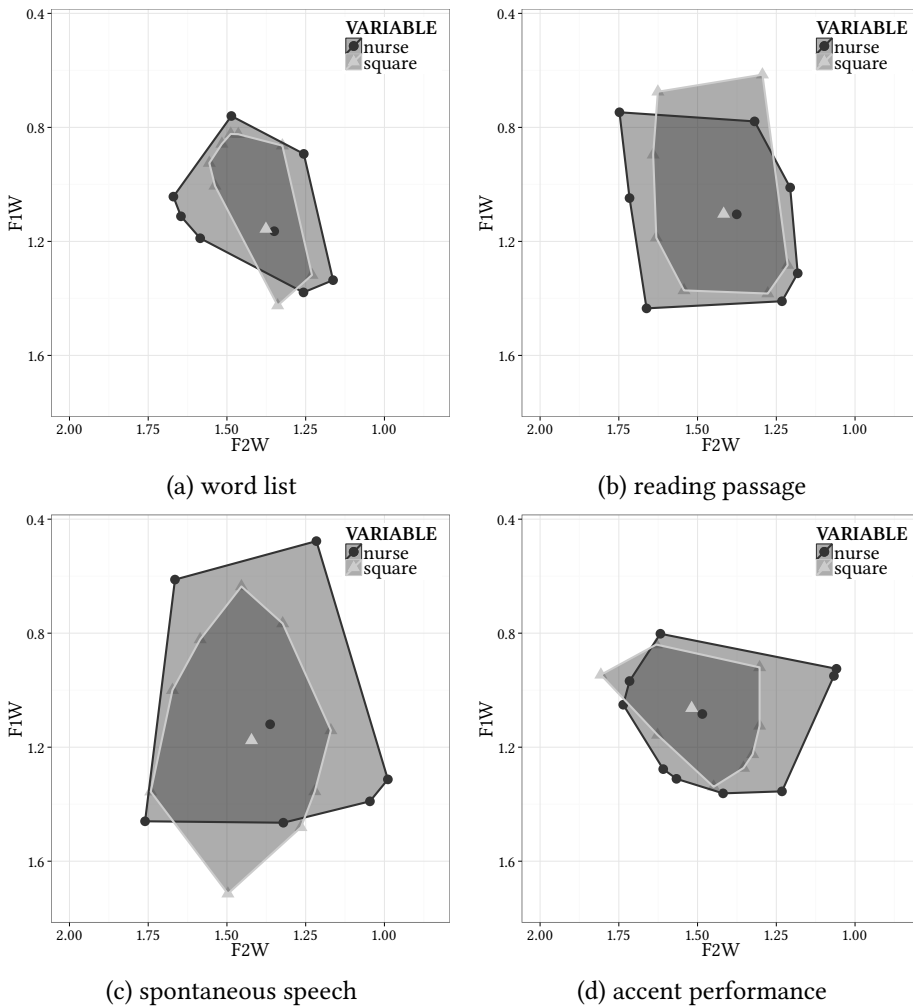


Figure 6.30: NURSE-SQUARE: vowel space by style (old speakers)

by the dark, that of *SQUARE* by the light polygons. The mean realisations are represented by circles (*NURSE*) and triangles (*SQUARE*) of the respective shades.

Two aspects strike the observer when inspecting the realisational spaces of these two vowels in the old generation: (1) The range of variation (as represented by the surface area of the polygons) increases considerably from the word list to free speech, via accent imitation and the reading passage. (2) The vertical extent of the polygons is much larger than the horizontal one, which means variants of both *NURSE* and *SQUARE* are relatively stable in the front-back dimension. It is possible that the large degree of homogeneity that is found for the word list is primarily due to the smaller number of observations in this sub-sample; if more data are collected (like for the free speech style in this study) the chance of including the occasional ‘extreme’ variant increases. It seems more likely, however, that subjects largely agree on which variants are appropriate for the more formal registers reading and word list (and, incidentally, also for the stereotype of Scouse!), but allow themselves to choose from a larger set of options in spontaneous speech. These options (in free speech and the other styles) seem to differ largely in F1, but a lot less so in F2. With the possible exception of Figure 6.30d vowel spaces are all larger in height than width, which indicates that the old speakers in my sample manipulate F1 quite considerably for these vowels, while producing all variants with a relatively stable F2.

When we look at the mean realisations of the two vowels (dark circle for *NURSE*, light triangle for *SQUARE* in the centre of the polygon) in this age group, we can, first of all, see that the distance between the centres of gravity is greatest in Figure 6.30c, which means that the higher Pillai score and lower p-value found for this sub-sample (cf. Table 6.16) is *not* only due to the larger number of observations in this style, after all. The distributions are indeed more distinct than in the other registers, if only slightly so. The relative positioning of mean vowel realisations is as would be expected: *NURSE* is both slightly higher and more central than *SQUARE*. Interestingly now, the mean of *NURSE* does not really move at all from spontaneous speech to the reading passage (Figure 6.30b). All the same, the average realisations of the two vowels actually become more similar (instead of more distinct, as would be expected for a more formal register), because *SQUARE* is both raised and centralised and thus approaches *NURSE*. For the word list (Figure 6.30a), both means are somewhat lowered and even closer to each other. This is, for *NURSE* at least, again contrary to expectations, since a move towards more standard realisations would involve raising rather than lowering. Only when we move from free speech to accent imitation (Figure 6.30c and Figure 6.30d), does the mean of *NURSE* change more drastically: it is considerably more front, and almost identical to *SQUARE*, which is fronted and raised as well.

6.2.3.3 Middle-aged speakers

When we turn to the data collected in the group of middle-aged speakers (Figure 6.31) and compare them to *NURSE* and *SQUARE* realisations of the old group, both similarities and differences emerge. For instance, *NURSE* is again always more variable than *SQUARE*, which is expected since *SQUARE* is supposed to be the steady target of the merger. Just as in the old group, there is also clearly

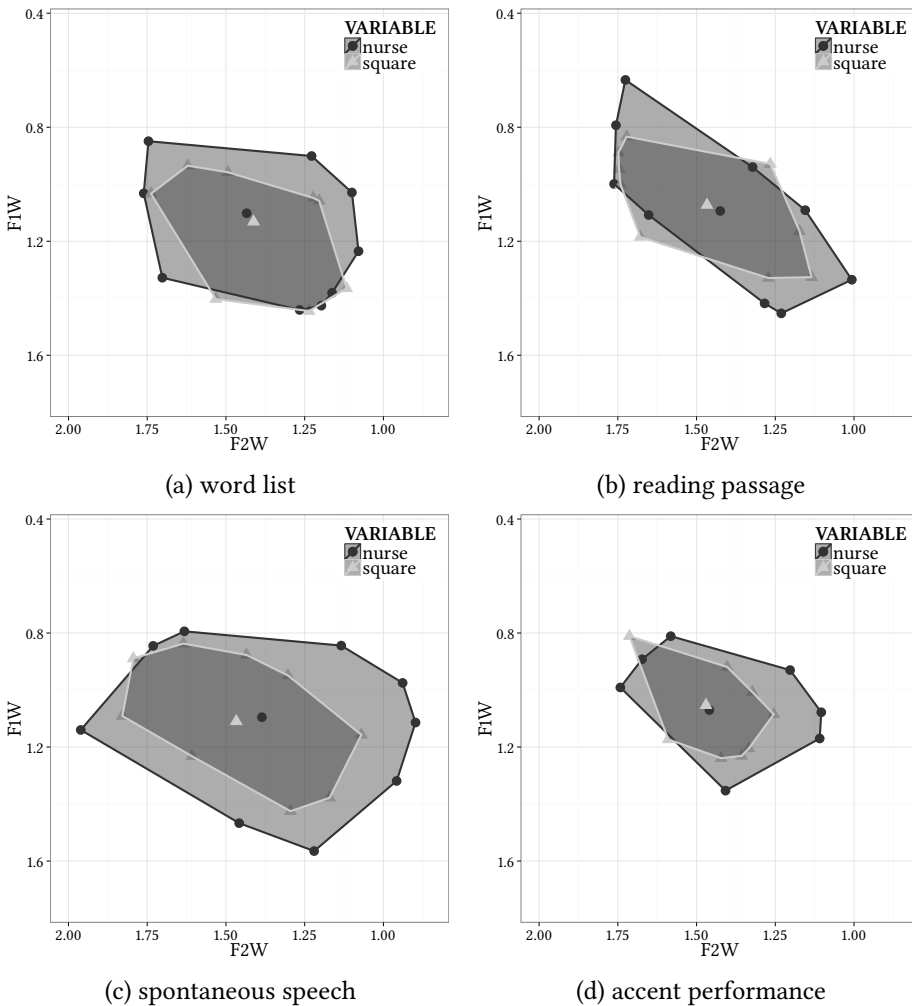


Figure 6.31: *NURSE-SQUARE*: vowel space by style (middle-aged speakers)

more variation (for both vowels) in spontaneous speech than in the other speaking styles. But in contrast to the older generation, the lowest degree of variation is found for accent imitation, while the word list and reading passage seem to be roughly comparable in this respect. However, it should be noted that the realisation spaces – particularly when comparing the word list, reading, and accent performance – differ less in area in the middle-aged group, generally, which is to say that the range of variation depends less on style than in the old group. All the same, it looks as if, again, the target realisations for the more formal and the stereotype contexts are clearer, or less controversial, than in spontaneous speech, where a wider range of phonetic variants is encountered. The most pronounced difference between the old and the middle-aged speakers is that in the former most variation is found in the height dimension, whereas in the latter group the horizontal extent of the individual vowel spaces is at least as great as, and frequently greater than, the vertical one, which means that the main axis of variation seems to be F2 in this age group.

With respect to the centres of gravity of these vowel clouds, there is something going on that is even more interesting than the pattern found for the older speakers. In spontaneous speech (Figure 6.31c), the relative positioning of NURSE and SQUARE is identical to the one found for the older speakers: NURSE is both slightly higher and backer than SQUARE, which corresponds to the setup expected on the basis of the standard if the vowels are not perfectly merged yet. The means of the two distributions are also further apart from each other than in the other styles, although the difference is smaller than in the old speakers. If this register is once again taken as the benchmark, or baseline, both NURSE and SQUARE change when people are asked to put on a particularly strong Scouse accent (Figure 6.31d): The former is fronted, the latter slightly raised. As a result, mean realisations of both vowels are essentially identical. When we go the other way from less formal free speech to more formal reading (Figure 6.31b), however, mean NURSE remains almost the same, if one ignores the tiny amount of fronting that is visible. In this register it is mostly SQUARE that moves: raising takes it closer to NURSE. This results in more *merged* distributions (cf. Table 6.16), instead of the more *distinct* ones that would normally be expected in a more formal speaking style. When subjects read out a word list (Figure 6.31a) SQUARE is centralised, while NURSE is fronted at the same time. The outcome of these processes is that NURSE (dark circle) actually ends up in a position that is *more front* than that of SQUARE (light triangle), thereby completely reversing the relative positioning found in spontaneous speech, at least with respect to the F2 dimension.

6.2.3.4 Young speakers

Figure 6.32 visualises the realisation spaces of **NURSE** and **SQUARE** for Liverpudlians aged between 19 and 29. These four sub-plots look clearly different from the corresponding ones in Figure 6.30 and Figure 6.31 because the areas of all polygons are considerably smaller. Smaller areas in the figure translate to a smaller range of phonetic variants that occur in this age group. The youngest speakers in my sample can therefore be said to be markedly more homogeneous in their

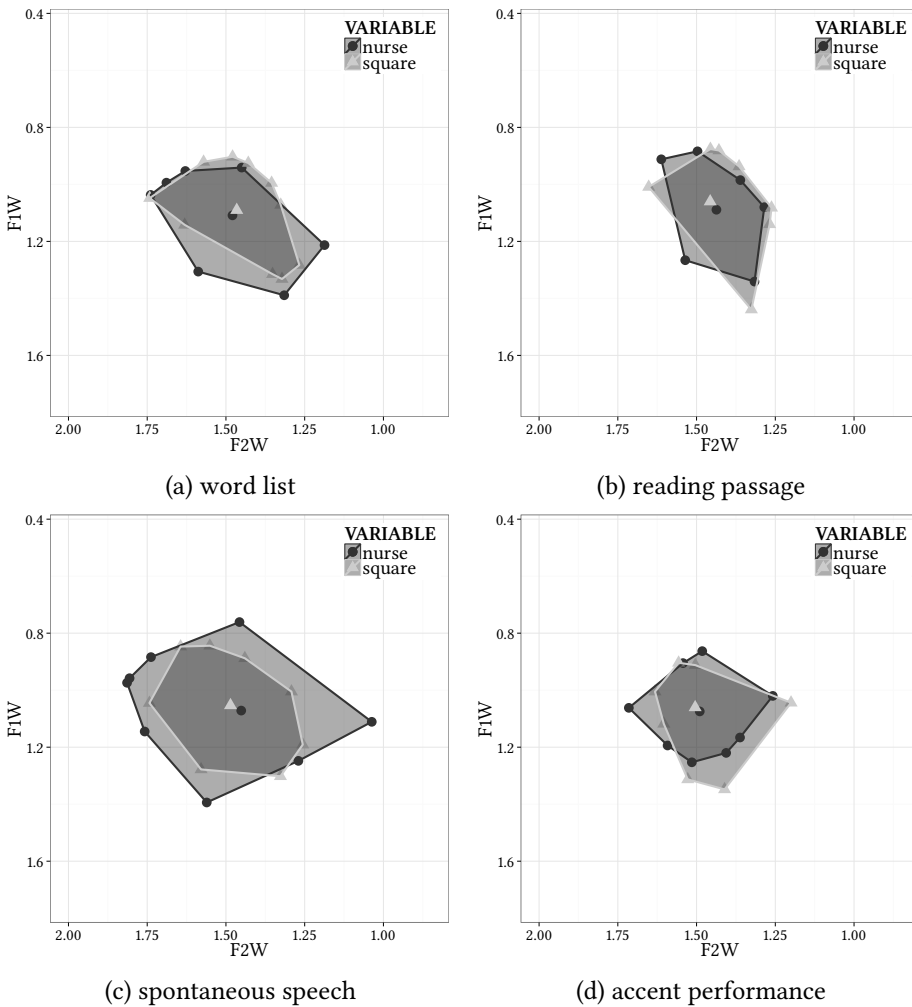


Figure 6.32: NURSE-SQUARE: vowel space by style (young speakers)

NURSE and SQUARE productions than speakers of their parents' or grandparents' generation. Style seems to play a much less important role in this respect than it did for the other two age groups. While it is true that there is more variation in free speech (Figure 6.32c) than in the other three registers, the difference is much less pronounced than for the middle-aged and the older speakers. The plots for the word list, the reading passage, and accent imitation, in turn, do not seem to differ at all in terms of polygon area. The youngest subjects thus seem to be much more agreed on target pronunciations in *all* styles, in contrast to the old and the middle-aged speakers, who show different levels of variation in different registers. Another difference with the young speakers is that, interestingly, NURSE is only considerably more variable than SQUARE in spontaneous speech, but not in the other three styles, where the realisational spaces of both vowels have roughly equal size. Both NURSE and SQUARE seem to have about the same level of stability.

Younger Scousers also set themselves apart when it comes to the mean values of these two variables. In spontaneous speech (Figure 6.32c), mean NURSE and SQUARE are closer together than in both the middle-aged and the old group, with NURSE being just a little bit more central than SQUARE and about the same height. During accent performance (Figure 6.32d), SQUARE essentially remains in the same location, while NURSE moves forward, but only a tiny little fraction (which is not surprising, given that it is already very front in free speech). Going from spontaneous speech to reading out a text (Figure 6.32b) neither NURSE nor SQUARE change very much, although the latter is slightly more central. When my youngest subjects read out a word list (Figure 6.32a) SQUARE is virtually identical (to its realisation while reading), but NURSE is marginally more front and lower to the same degree. In summary, one can say that the youngest speakers in my sample do not only show less variation *within* a particular speech style, but they also show next to no change *between* styles.

6.2.3.5 Age means and individual differences

Having investigated style differences within the age groups on the previous pages, it is now time to take a step back again and briefly focus on the overall effect of age on the realisations of NURSE and SQUARE. Figure 6.33 shows the mean realisations of both vowels for all three age groups. The means plotted in Figure 6.33 summarise across the style dimension (i.e. one grand mean was calculated per age group), but of course these means are biased towards spontaneous speech realisations as they account for the large majority of observations.

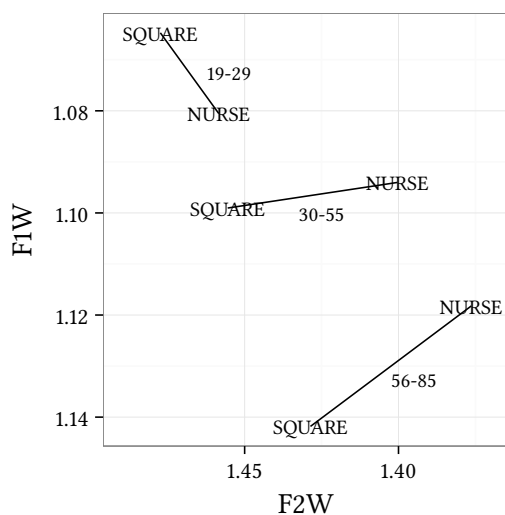


Figure 6.33: NURSE-SQUARE: mean vowel position by age

This graph essentially tells us three things: (1) In all three age groups, NURSE is (still) more central – to the right in the graph – than SQUARE. (2) The means for the old speakers are to be found in the lower right corner, the ones for the middle-aged speakers in the middle, and the averages for the young group lie in the upper left corner, which means that both NURSE and SQUARE simultaneously become fronter and higher from the oldest to the youngest speakers in my sample. While it has to be said that these differences are subtle (note the scale of the graph), they have been shown to be statistically robust in §6.2.1 and §6.2.2. (3) The decreasing distance between NURSE and SQUARE from the oldest to the youngest subjects signals that the merger becomes progressively more complete in my sample.

While a detailed analysis of individuals is beyond the scope of this study, it is important to note that all figures and tests in this section which generalise across age groups often mask considerable differences between the members of these groups. This is obvious from Figure 6.34, which plots Pillai scores of individual speakers (across all styles) on the y-, and their age on the x-axis (age groups delimited by vertical lines). The third variable represented in this graph is social class of the speaker, because it was found to be part of significant three-way interactions with style and age in §6.2.1 and §6.2.2: Pillai scores of middle-class speakers appear as black circles, those of working-class subjects as light grey ones.

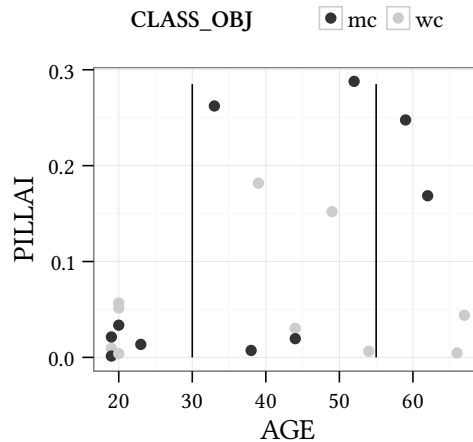


Figure 6.34: NURSE: Pillai scores by age and social class (individuals)

This setup serves to illustrate an interesting relationship between Pillai scores, age, and class. Towards the right-hand side of Figure 6.34 one can see that old working-class speakers have much lower Pillai scores than middle-class subjects of the same age, which is not surprising, given that having the merger (and thus a low Pillai score) is a non-standard feature. However, in the middle-aged group this distinction has already become a lot more fuzzy: one half of the participants in this sub-sample has Pillai scores near 0, the other half produces scores that are considerably higher (while still comparatively low in absolute terms). Crucially, these scores no longer correlate with class, because both the merged and the distinct sub-group include both middle *and* working-class speakers. In the youngest group, finally, there are only merged speakers, and social class is therefore no longer a relevant factor. Only this last age group therefore really acts like a group. The middle-aged and the old group, in contrast, are characterised by a considerable degree of individual and class-related differences between subjects.

Sound change, priming, salience

This volume investigates the realisation and perception of four phonological variables in Liverpool English (Scouse), with a special focus on their sociolinguistic salience. Younger speakers' speech is found to be more local, but only for the two salient variables in the sample (NURSE-SQUARE and /k/ lenition), which appear to carry considerable amounts of covert prestige. Local variants of non-salient happy-tensing and velar nasal plus, on the other hand, are actually found to be receding, so at least to a certain extent Scouse also seems to be participating in regional dialect levelling.

The importance of salience is also obvious in the perception data, with only the two highly salient stereotypes generating robust effects in a social priming experiment (albeit in the unexpected direction). These results indicate that the investigated variables differ measurably not only in their use in production, but also in terms of how central they are to mental sociolinguistic representations of Scouse. They also tell us more about the way we process, store, and (re-)use sociolinguistic variation in perception. By defining likely contexts for significant priming effects they might finally even help in coming up with a more elaborate 'theory of priming' in the realm of sociophonetics.

ISBN 978-3-96110-119-1



9 783961 101191