

NOTES ON THE EXERCISES

CHAPTER 1

Comments exercise 1

We have highlighted below the specific components of the introductions that identify how the writer found and justified their research topic. After each extract we have also given a synopsis of the main issues.

Rajah-Carrim, Aaliya. 2007. Mauritian Creole and language attitudes in the education system of multi-ethnic and multilingual Mauritius. *Journal of Multilingual and Multicultural Development* 28: 51-71.

On creole-speaking multilingual Mauritius, languages act as important markers of identity (Eriksen, 1998; Stein, 1982). In fact, most of the 12 languages present on the island are associated with specific ethnic and/or religious groups. The various languages can broadly be divided into three groups: ancestral languages (Indian and Chinese languages) whose usage is limited, colonial languages (English and French) and language of everyday interactions (Mauritian Creole/Kreol – see Note 1) (Rajah-Carrim, 2005). While most of these languages have a place in the education sector – as medium of instruction or subject – the native language of most Mauritians, Kreol, tends to be excluded from the classroom.

The teaching of languages has become a highly politicised issue in Mauritius. In 2004, the Minister of Education declared that Kreol would be officially introduced in the education system in the coming years. The new political leaders who came to power in 2005 are also committed to the promotion of Kreol. But how do the Mauritians themselves feel about the introduction of Kreol in the school system? In this paper, I discuss attitudes to the use of Kreol in the education sector based on a survey conducted in Mauritius. In the first section, I describe the demographic and linguistic situations of Mauritius. This is followed by a description of the national education system. I then show how the language-in-education issue is not unique to Mauritius and is tied to ideologies of identity and power. In the fifth section, I discuss the questions related to the school domain in the survey. In the next two sections, I analyse responses to the survey questions. The final section consists of a brief summary and conclusion.

Research topic set-up:

Mauritius is multilingual; language is an important marker of identity. Kreol may be introduced in schools, but how do people feel about this?

Research focus:

Change is taking place in Mauritius schools. Ideologies about language use in education are tied to issues of power and should be investigated widely.

Stuart-Smith, Jane, Claire Timmins and Fiona Tweedie. 2007. 'Talkin' Jockney? Variation and change in Glaswegian accent. *Journal of Sociolinguistics* 11: 221-260.

[...] In the late 1990s, preliminary results from a study of Glaswegian accent indicated that working-class adolescents, with few apparent opportunities for contact outside the city, were both using features usually associated with southern English (e.g. TH-fronting, the use of [f] for /θ/ in e.g. *think*), and at the same time, not showing expected 'Scottish' features (e.g. production of postvocalic /r/ in e.g. *car*). These findings led to a flurry of media reports which jokingly dubbed the 'new' dialect 'Jockney' ('Jock' = Scot + 'Cockney'), and which speculated that another possible cause for such patterns of variation was watching London-based TV soap operas, such as the popular show, *EastEnders*.

This paper presents the first integrated account of the linguistic facts behind such reports. Here we describe the main patterns of variation observed in a socially stratified sample of Glaswegian collected in 1997, and we account for them in terms of evidence which directly relates to the corpus itself, the recent social history of Glasgow, and indications of local language ideologies constructed about this history. Using univariate and multivariate statistical analysis, we look at the use of a range of consonant variables in speakers from two neighbouring areas of Glasgow, representing distinct points on the sociolinguistic continuum of Glaswegian English. [...]

Our young innovators are using a consonantal system which in many respects is more similar to that of London English, but at the same time, they are exploiting 'non-local' variation in such a way that it is used, and feels to them to be, thoroughly local. This then provides further support for the notion that constructing identity through linguistic variation is crucially connected to the local context (e.g. Labov 1963; Eckert 2000; Dyer 2002; see also Meyerhoff and Niedzielski 2003).

Research topic set-up:

Changes in Glasgow English seems to show influence of southern British English.

Research focus:

They look at a range of variables to get a broad picture. They evaluate the emergence of new forms as new forms of expressing localness.

Sharma, Devyani and Lavanya Sankaran. 2011. Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change* 23: 399-428.

In situations of migration, severe disparities between parent and peer dialects can arise for local-born individuals, who may have parents who are non-native speakers as well as peers who are native speakers of the local language. For such cases, Chambers (2002) proposes a strong peer-orientation mechanism: an innate accent filter that blocks parental non-native features and leads local-born children to exclusively acquire the local dialect. However, numerous studies have found that foreign phonetic features introduced via in-migration are not always lost in local-born speech (Sankoff, 2002). A "weak" view of dialect assimilation in migration might propose that, rather than being entirely lost, foreign traits can be retained and functionally reallocated.

Both the strong and weak views often treat nativeness as a major boundary. Accent traits are expected to be either absent (strong version) or immediately reallocated by the first set of individuals to acquire the local dialect natively (weak version). Incremental stages of intergenerational accent change have not been studied closely enough in immigrant groups to move beyond speculation, however, and the following question remains largely unresolved: How quickly and how completely do local-born generations acquire a local dialect and lose exogenous traits, and is this rate and degree of shift governed by largely cognitive (e.g., nativeness) or social (e.g., demographic) factors?

Research topic set-up:

Strong and weak views of dialect assimilation make contrasting predictions.

Incremental stages of intergenerational accent in immigrant groups need closer attention.

Research focus:

Previous work has not looked at the process by which children of migrants pick up local speech norms so we don't know whether what drives this is primarily social or cognitive.

Comments exercise 2

The following features appear to vary in these excerpts:

(ing) is used variably as *-in* and *-ing*
(to) is used variably as *to* and *tae*
Gonna is used variably as *gonna* and *gonnae*
I is used variably as *I* and *ah*

There is also one occurrence of what appears to be the discourse particle *like*, which varies between *like* and a zero realisation. Some other vernacular features are used but they do not vary in these transcripts: *guid* and *aye*. There are also several potential phonetic variables (e.g. (t)-glottalling), but they are not visible in this transcript.

Choose (ing) as a variable for further investigation. This is interesting because *-in* is used very frequently in Edinburgh conversational speech. The current status of (ing) in Edinburgh is to a large extent a continued reflection of the historical development of (ing): the spread of using [ɪŋ] for variants of the apical form. The variant [ɪŋ] in participles came to Edinburgh as an external norm. In most Middle English varieties, the participle and verbal noun forms had been conflated under *-in/-ing* (Görlach, 2002:96). In Middle Scots, participles remained distinct in the written language until the 16th century, and *-and* and *-yng* were clearly distinguished. Görlach argues that even after that, in spoken Scots the two forms remained distinct for quite some time as *-and* forms were rarely used wrongly for the verbal noun. They eventually fell together as [ən] or [ɪn]. This pronunciation merger and the formal similarity of *-yng* and *-ing*, Görlach (2002:96) argues, resulted in the take-over of *-ing* for the participle. Although for quite some time *-and* continued to have an identificational function in some literary Scots before it fell completely out of use in the late 20th century.

As for constraints, Meyerhoff and Schlee (2013) investigate the following constraints for (ing) in the speech of Edinburgh adolescents:

1. Grammatical category: verb; noun; proper noun (e.g. *Reading, Carling*); pronoun (e.g. *something, nothing*); discourse marker (e.g. *or something*); deverbal adjective; gerund; preposition (*during*)
2. Preceding context: apical (e.g. *writing*); velar/glottal (e.g. *talkking, nothing [nɔhɪn]); other*
3. Following context: apical (e.g. *sending letters*); velar/glottal (e.g. *sending good wishes; doing Higher Geography*); other; pause
4. Preceding alveolar/velar nasal in the word: alveolar nasal; velar nasal; neither
5. Realisation of previous (ing) variable: [ɪn]; [ɪŋ]; no previous (ing)
6. Number of syllables in the word: 1; 2; 3, 4, 5
7. Word frequency (as LOG(10)+1 in the British National Corpus)
8. Style: conversation; reading
9. Sex: female; male
10. Age

Findings could be relevant for linguistic theory in regards to spread and status of *-ing* in Scotland and its influence on language change and the establishment of

community norms. If a variety of frequency measures are included in the analysis, results could be made relevant to an evaluation of usage-based approaches to language change.

For those interested in qualitative data analysis, the recordings offer good data to explore style shifting or the expression of different identities and how the speech acts that are being performed influence the language that is used. In the first excerpt the coach gives directions, in the second excerpt he discusses availability for the next game and in the third excerpt the coach gives advice. Note how the third excerpt includes a lot of *ah*, and that it is here where the only occurrence of *-ing* appears. The pragmatic functions of some of the linguistic structures, e.g. *okay*, could also be explored qualitatively.

References

- Görlach, Manfred. 2002. *A textual history of Scots*. Heidelberg: Universitätsverlag C. Winter.
- Meyerhoff, Miriam and Erik Schlee. 2013. Hitting an Edinburgh target: Immigrant adolescents' acquisition of variation in Edinburgh English. In Robert Lawson (ed.), *Sociolinguistic perspectives on Scotland*. Basingstoke: Palgrave, 103-128.

Comments exercise 3

Again, we have highlighted the segments of the conclusions that seem to us to identify further research topics. We synthesise them and suggest what might be needed to pursue them.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453-476.

Conclusions

Labov's introduction of class into the study of language was a landmark of immense importance. But his view of the class hierarchy, and of the relation of standard and vernacular language to that hierarchy, is only the beginning of a theory of the social value of variation. The social is not just a set of *constraints* on variation – it is not simply a set of categories that determine what variants a speaker will use – it is a meaning making enterprise. And while one's place in the political economy has an important constraining effect on how one makes meaning, and on the kinds of meanings one engages with, this place cannot be defined in terms of a simple model. A theory of variation ultimately must deal with meaning, and not only does a view of meaning in variation as predetermined and static seriously undershoot human capacity, it cannot even account in any principled way for the changes in correlations that have been observed over the lifetime of a sound change (e.g. Labov 2001: chapter 9). Ultimately, all change unfolds in the course of day-to-day exchange, and that exchange involves constant local reinterpretation and repositioning. Ultimately, it is in this action that we can get at the meaning-making that gives life to variation. While the larger patterns of variation can profitably be seen in terms of a static social landscape, this is only a distant reflection of what is happening moment to moment on the ground.

Areas for further investigation: Detailed analysis of how variation impacts on meaning in face to face interaction.

What you will need: Analysis of the larger patterns of variation in a community and detailed, repeated interactions of individual speakers.

Llamas, Carmen, Dominic Watt and Daniel Ezra Johnson. 2009. Linguistic accommodation and the salience of national identity markers in a border town. *Journal of Language and Social Psychology*

28: 381-407.

Conclusion

This article has presented the results of a study designed to test the extent of speakers' linguistic accommodation to members of putative in-groups and out-groups in a border locality, where such categorizations can be said to be particularly accentuated. Findings were considered in terms of their implications for the notion of salience, the evidence required for claims of phonological convergence and divergence, and the interviewer effect in the compilation of data sets for use in quantitative studies of phonological variation and change. [...]

As regards compelling evidence for phonological divergence in intergroup short term contact interactions, it appeared that to interpret patterns of variation in such a way, it would be necessary to identify a set of default production patterns, presumably those corresponding to a vernacular speech style, from which the speaker would move. It is not clear that this identification is possible. [...]

More work is necessary on the central notions of interspeaker convergence and divergence in short-term contact situations, which are so often invoked in the interpretation of patterns of phonological variation and change. The potential they contain as explanations of motivations for variable linguistic behavior implies that better understanding of such processes permits us to comprehend more fully the dynamics of language change more generally.

Areas for further investigation: How do speakers converge and diverge in short-term contact situations? What are individual speakers' default patterns of variation? Are these associated with their most vernacular speech style? What you will need: Detailed case studies of individuals who speak different languages or dialects who have only short or fleeting contact with each other.

Comments exercise 4

Select a set of skills which you think will help you further your career and incorporate these into your project. For example, leadership and teamwork skills can be honed in group projects. If you want to develop quantitative or qualitative skills, choose specific methods of data collection and analysis. Mastering information communications skills would require you to choose a project that involves such skills as online survey design for a questionnaire-based study or video recording and analysis for an ethnographically-based study.

CHAPTER 2

Comments exercise 1

The variable is third person singular subjects.

The variants are NPs, demonstratives and zero subjects. Later Skarabela et al. amend this and note that there are two kinds of demonstrative – full forms and clitic forms. This shows that new variants may become apparent as part of the research process.

Comments exercise 2

We can make only very tentative generalisations based on these three excerpts. We would certainly have to consider more data to fully answer these questions. First, *with* does not appear to be variable in these excerpts. It is always *wi*.

To is variable as *to* or *tae*. We would want to consider a variety of contextual factors to delimit the envelope of variation. One such factor could be word type. *To* appears as a preposition (going to), in infinitive forms (to combat, to ram) and in *into*. *Tae* also occurs in all of these, so they all form the context of variation for this variable. *Tonight* is always realised as *tonight* and may not be part of the context of variation, and its lexical status is another reason to exclude it. We would have to have a closer look at expressions such as *used to*, *have to* and *able to*. They appear to be mostly used with *tae* and may not form part of the variable context. *Supposed to* is realised with *to*, so this may be an indication that this hypothesis may be wrong or only partially correct.

CHAPTER 3

Comments exercise

Nicole has a problem. People might question whether she is exploiting her position in the day care to her advantage. The kids, the kids' parents and her boss haven't been informed about what she's doing and what she plans to do with the data she collects. On the other hand, there is nothing to stop Nicole observing what's going on around her. That's how much good research starts. If she gets permission now, she might be able to use her notes as background information even if she can't incorporate them into her analysis.

Karen has up to this point no informed consent to record residents in the nursing home. The nurse is not in a position to give such consent, and it also appears that the nurse is not fully aware what Karen plans to do. This is even more so the case for the residents in the nursing home, some of which may be classified as vulnerable adults. It must be Karen's priority to obtain research permission from her institution and consent from individual residents. An appropriate procedure may be to ask the nurse to introduce Karen to some of the residents, and Karen can then ask them whether they'd be interested in an informal chat with her, during which she can explain her project and obtain informed consent in writing from individual residents. She can then schedule her interview for another date. Karen must ensure that everyone understands what is going on, and that residents have volunteered to participate and be recorded.

Russell intends to make covert recordings. The respondents to his question would not be aware that they are being recorded, so this is not an ethical approach to take. There are published studies in which people were approached in a public setting but no recording was made. In these studies, the response was short and precisely targeted, so that the researcher could remember and note down what was said. Sometimes, a second researcher may be involved whose job it is to memorise what the person said. But even in these scenarios informed consent to use the data was not obtained and University Ethics Boards would probably not approve this procedure. This is not publically available data and should not be treated as such. The researcher has specifically initiated a conversation pretending to need help, and respondents were kind enough to be helpful, so this conversation would not otherwise have occurred. There may be better options to find out about this particular linguistic issue. These may include role plays, discourse completion tests, or recording such interactions in a space that potential participants enter after being informed that an interaction may occur and that this interaction could be recorded and be used for research purposes.

CHAPTER 4

Comments exercise 1

1. Random sampling, stratified random sampling or judgement sampling
2. Depending on the specific question: judgement sampling, social networks and the 'friend of a friend' technique, communities of practice and ethnographic methods
3. You could focus on this specific context and conduct a social network or community of practice study, combined with ethnographic methods. Alternatively, you could make use of this context to start a larger quantitative survey study, using judgement sampling, which you then extend to other contexts.

Comments exercise 2

A focus on a small community of speakers makes it suitable to students who work on a small research project. For example, it may be worth investigating how large-scale sociolinguistic patterns in a city – see Cheshire, Fox, Kerswill and Torgersen (2008) for London and Labov, Rosenfelder and Fruehwald (2013) for Philadelphia – play out in a small group of people and how these linguistic features are used as resources in the construction of identities. If, after having read a thorough quantitative study, you know that in your city of interest women are leading a change or that a linguistic feature is stratified by class, you may want to look at how these features are used or avoided within a small group of people. We give some advice on how to access communities in chapter 4.

Such a study would be important because it allows you to investigate a set of linguistic features in a small community that is likely quite homogenous in its demographic characteristics. You can study how precisely a feature is used in interaction and what interactive work it does. You can find out about its social meaning, and by doing this, shed more light on its sociolinguistic status as you situate your findings in the bigger picture that you have previously read about.

Comments exercise 3

Schleef et al. (2011) accessed the schools by contacting head-teachers and asking for their collaboration in the project. A teacher was normally assigned to the project, who then helped recruit students for interviewing. Schleef et al. (2011) used judgement sampling because they were interested in a very specific group of speakers, adolescents from Poland, rather than the population of London or Edinburgh as a whole. They were interested in finding out whether, when and how Polish adolescents acquire English vernacular features, which made a quantitative approach more suitable and judgement sampling the best way to collect the data. Once they had identified their population, they collected data from a number of speakers paying attention they get a mix of males and females and different age groups. Social background was noted as well. However, social class is difficult to consider in a school-based study as students often come from the same neighbourhood, which may be class-specific. In addition, the focus on Polish adolescents made the concept of social class very difficult to deal with. Immigrant populations and adolescents are notoriously difficult to assign to a specific class, because class assignment is most likely based on the class of their

caregivers, who may have worked in a traditional middle class job in Poland but work in a job that many would regard as working class in the UK. In addition, caregivers of a child may very well come from different social classes.

Another limitations in collecting data from Polish adolescents was to find a sufficient number of students willing to be interviewed. The number of Polish adolescents in each school was very limited, so the sample Schleef et al. (2011) could compile was limited as well. Data from English/Scottish students was collected because Schleef et al. (2011) had to find out what the local vernacular norms were among adolescents at the respective school in London and Edinburgh. Without this information, it would have been very difficult to make any claims about the acquisition of these norms by Polish adolescents.

CHAPTER 5

Comments exercise 1

These modules could relate to local history and customs, local discourses, local or current events in the region. Before you interview people, you should always make sure you know what people like to talk about and how they talk about it, you want to know the local sports teams (and, potentially, their ethnic or religious affiliations), what people do for fun, what places they go to, etc. Have a look at some of the modules on our companion website, pick a relevant area and design your own modules.

Comments exercise 2

The problematic question here is “Oh, because it’s not right on Malo”. Malo is the island they live on. Yes/no-questions can be answered with a simple *yes* or *no* and do not serve the purpose of keeping the interviewee talking very well. Miriam realises the effect of her question right away and turns it into a wh-question, which elicits more information and directs the interviewee back into her narrative. Apart from wh-questions, other strategies which ensure that the interviewee, rather than the interviewer, talks are questions for explanation or elaboration: “can you tell me more about...”, “what happened then...”, etc.

Comments exercise 3

The decision tree consists of eight contextual criteria originally designed to facilitate the analysis of spontaneous speech and assign stretches of speech to one of these eight contexts. The further down you go the tree, the more difficult it is to assign language to one of the contexts. Consult Labov (2001) to find out what the various categories mean.

However, this decision tree also helps us structure and plan our interviews. The question we are asking is about transitioning between different styles and topics. How can this best be done? You may want to aim for careful speech first because, most likely, this is the type of speech that interviewees will be using at the beginning of the interview anyway. So question-answer responses on topics are a good start but you will have to move past these if you also want to record longer stretches of casual speech. Similarly, if you’ve reached a point in an interview where the conversation is relatively casual, it might be difficult to “pop” back up to a topic on the other side of the tree in a naturalistic and unselfconscious way. Preparation is the key here. You want to prepare detailed interview modules that help you move from topic to topic in order for you to record sufficient amounts of speech in different styles.

You may also want to consider the extent to which those topics necessarily line up as formal/informal. Imagine someone was telling you about school games or memories from when they were a child, and they start spontaneously talking about features of language that differentiate their childhood from younger people’s nowadays. For example, Miriam found in her Bequia interviews that when people were asked to talk about what it was like growing up on the island when they were little, some people told her how in the old days you had to say hello to every adult you passed on the road and if you didn’t you would get hit

(often twice, once by the adult you had disrespected and once by your parents when they heard about it), whereas nowadays, kids are seen as showing little respect to them. These kinds of events are hard to categorise within the careful/casual dichotomy that is set up here.

References

Labov, William. 2001. The anatomy of style-shifting. In Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, 85-108.

CHAPTER 6

Comments exercise 1

First, you observe. Since you'll do your observation in a coffee shop, you will most likely have bought a beverage, i.e. you're already participating as well. You can find out a lot by observation alone. If you position yourself well in the coffee shop, you may also be able to overhear the service encounter, and you may be able to focus on language right from the start. One potential focus of your study could be the language formulae used to order a coffee: *can I have, could I get, I need, a coffee please*. A sociolinguist may be interested in finding out whether there is any link between how these requests are made and social categories. These could be global categories such as male versus female, but they may also be more local categories that you may be able to uncover through observation. You will take copious notes, writing down the setting, the participants, clothing, who speaks, who is silent, etc.

If observation on its own is not enough, you could also conduct interviews or short surveys to find out more about staff and customers. You may also record some interactions. This has the advantage that you could go back to the coffee shop, play recordings to staff and ask them for their perspective on the activity observed. You may also be able to extend participation and work in the coffee shop for some time in order to get the perspective of both staff and customer. Then pull all the different data, ideas and arguments together: how does the language used during service encounters, the social organisation of the encounter, how different staff and customers see themselves and the other in these encounters come together to create a unique cultural activity that is characteristic of the people involved?

Why might this be important? Apart from potential descriptive and theoretical aspects, there is one very obvious applied aspect. These kinds of encounters are to an increasing degree intercultural encounters. Understanding how staff and customers see themselves, but, more importantly, what their norms and expectations are can be immensely helpful in intercultural training. John Gumperz is known for conducting a study at a cafeteria at Heathrow airport. It involves British employees eating at the cafeteria and Pakistani and Indian cafeteria workers. The cafeteria staff were accused of rudeness by the British employees and after many angry words had been exchanged, John Gumperz stepped in to find out what was going on. He noticed that the main source of the conflict was the way the cafeteria staff offered gravy for the food. While British English speakers would expect rising intonation, *gravy?*, the cafeteria workers used their own cultural convention of offering the gravy with falling intonation. Although, they had polite intentions, the falling intonation was taken as a contextualisation cue indexing impoliteness.

You may very well make similar discoveries in your own study.

Comments exercise 2

(1) Interview data would be extremely difficult to collect from children who are often reluctant to speak to strangers. In addition as Smith et al. (2007) point out,

a lot of data is required from children as they normally do not speak continuously for an hour.

(2) The participant selection process was very specific. They asked participants to self-record.

(3) All issues discussed under “Recording Spontaneous Speech” are relevant for their study: The observer’s paradox is exacerbated in the case of young children, thus self-recordings were conducted. The amount of speech one obtains from various people differs. It is unclear how this was addressed but since only 5 hours were transcribed per dyad, this will have mitigated the issue. If context is not specified, the recordings one receives may have been made in a very wide range of settings and may not allow direct comparison. This may have been less of a problem for caregiver/child dyads as there may have been less variation in the context where interaction takes place. There are further issues to consider when recording spontaneous speech. The researcher may be present as an imagined audience, which may have been an issue here. Further, when recording spontaneous speech, the quality of recordings may be poor, in terms of background noise, recording volume or even the type of speech recorded. However, the researchers provided their own recording equipment and report that the recording quality was sufficient.

(4) The corpora are generally comparable, although it is clear that they differ. The caregiver/child corpus is based on naturally-occurring language in a range of contexts while the adult-to-adult corpus is based on interview speech only. This must be borne in mind when conducting the analysis. However, the authors point out that the style in the interviews is very informal. In addition, it must be kept in mind that even if the adult-to-adult corpus had been based on naturally-occurring language, it would still be different from the caregiver/child corpus, as the latter consists mostly of child-directed speech, which is very different from adult-to-adult speech anyway. The current adult-to-adult corpus will still serve its main purpose, namely to represent the norms of usage in everyday conversation to which caregiver/child interaction can be compared.

References

Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.

CHAPTER 7

Comments exercise 1

(1) You'll need historical corpora to study this change, which Kytö (1993) found to take place over the course of the 17th century. A corpus that is annotated for part of speech will speed your work along. This is because searching your transcripts simply for -s will find plural markers, which are not relevant to your study, so being able to restrict your search just to third singular present -s will eliminate these false positives.

(2) This is a variable you could study either in speech or in writing (and it could be interesting to do both and compare the two). We can't think of anything you would gain by studying this variable in an audio corpus, as the variation strikes us as purely morphosyntactic, so transcriptions would suffice for your speech component. This study would be most efficiently done in a corpus that is syntactically parsed, to allow you to quickly eliminate instances of 's and *of* that are not part of the envelope of variation (e.g. *John's my best friend, out of the window, a loaf of bread*).

(3) You'll need an audio corpus to study this variable: even a phonetically transcribed corpus won't mark degrees of vowel frontedness. You may want to consider whether vowel stress plays a role in the variation, and whether you could easily obtain this data. Software for forced alignment and vowel extraction (see Breakout Box) is likely to provide it for you, and has the added bonus of saving you time by coding your data for you. You will also want access to information on speakers' geographical region, as not all regional varieties show this variation, and perhaps also their age and sex, as this is a change in progress in many varieties.

(4) This variable is not likely to be accurately represented in orthography, as transcribers are often instructed to correct non-standard contractions (and, indeed, our spellchecker flagged up *woulda* as we were writing this). We thus recommend studying a variable like this in an audio corpus, where you can judge for yourself which variant is used.

Comments exercise 2

There are a number of ways contextual style could be studied in a present-day corpus. Some corpora have been explicitly designed to contain language from different registers, like the Corpus of Contemporary American English, which contains spoken, fiction, newspaper, magazine, and academic genres, or the Santa Barbara Corpus of Spoken American English, which, alongside face-to-face conversations, contains language used in more formal contexts, like meetings and lectures. If the corpus you're working with hasn't been designed for the examination of contextual style, you may be able to come up with some proxies. For instance, comparing conversations between strangers (like those found in the Switchboard and Fisher corpora) to conversations between friends (like those found in the CALLHOME and CALLFRIEND corpora) could give you a reasonable approximation of style. In a corpus of sociolinguistic interviews, you could compare speech from the beginning of the interview to speech from the

end, under the assumption that as the interview subject becomes more comfortable, his/her speech becomes more casual.

Researchers have found various ways of approximating style in historical corpora. A corpus of correspondence can allow you to use the social distance between the writer and the recipient as a proxy for style. This is the approach taken by D'Arcy et al. (2013), who distinguish letters to nuclear family, letters to friends and non-nuclear family, and letters to business associates. Warner (2005) uses a measure of lexical complexity—a composite of type–token ratio and average word length—as a proxy for style in his corpus of historical English texts. The work of Douglas Biber (e.g. his 1988 book) provides extensive discussion of register correlates in written language.

References

- Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- D'Arcy, Alexandra, Bill Haddican, Hazel Richards, Sali A. Tagliamonte, and Ann Taylor. 2013. Asymmetrical trajectories: The past and present of *-body/-one*. *Language Variation and Change* 25: 287–310.
- Kytö, Merja. 1993. Third-person present singular verb inflection in early British and American English. *Language Variation and Change* 5: 113–139.
- Warner, Anthony R. 2005. Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change* 17: 257–280.

CHAPTER 8

Comments exercise 1

You can verify participants' demographic characteristics in a face-to-face setup. It also allows you to audio record responses, if that's likely to be of help, and to ask clarification questions based on people's initial responses.

Long-distance surveys on the other hand allow us to collect more data faster. They come in various forms: they can be sent by e-mail or distributed by hand or on the internet. Telephone, computer or internet surveys make it possible to play recordings even to participants who are miles away. Online surveys also afford interesting options for responses (e.g. asking people to use moving sliders, see example in Chapter 9), and they can be distributed widely fast. But be careful! These surveys attract many dishonest responses, especially when participation results in some kind of reimbursement, such as a voucher. There are tricks for identifying chancers who are trying to game the system. Web-based survey sites normally provide you with information so you can remove participants who took the survey in an unusually short period of time, who were not in the country they claimed they were when they took the survey or those who always checked the same box.

Always consider who may have filled in your questionnaire. Was it based on self-selection? A particular medium? This may bias your findings. For example, only people with internet access can fill in online surveys, and they are more likely of a particular socioeconomic class or follow a particular life style.

Comments exercise 2

There is a whole host of terms that can be used to describe stimuli. Campbell-Kibler could not include all of them. She had to eliminate some. Most likely, these decisions were based on what terms respondents had reacted to in her previous research, and she will likely also have avoided using highly similar terms. She also selected some question types that take up little space. She starts with semantic-differential scales, and then moves on to rating scales and gives one multiple-choice question on speaker age. These take up quite a bit of space. Rather than continuing with rating or semantic -differential questions for words such as *confident*, *reliable*, etc., she has chosen to use checklists for these instead. This allows her to accommodate a larger number of words in less space. The downside is, of course, that participants may attend to these less than to semantic-differential tasks and rating scales. She closes with one open question, which is always a good idea to include. It does not take up much space but some respondents may not fill it in (properly) at all. These, nonetheless, give you some very nice bits of qualitative information about the voices heard. This survey features a very good selection of question types. Some survey programmes also allow you to use a slider rather than check box scales, which some respondents may enjoy using.

Different statistical procedures have to be selected for the question types used. If respondents answered on a scale (e.g. for the semantic-differential and rating scales), the researcher will be able to quantify these adjectives, e.g. *intelligent*,

with values ranging from 1 to 6 (i.e. the dependent variable is continuous). So you end up with 6 potential answers that may average out across groups and give you a score of, let's say 2.5 for all males and 3.0 for all females. If you think you can master inferential statistics, linear regression is often used to find out which scales are significantly different for the variants heard. If respondents gave a binary answer (yes, no: i.e. the dependent variable is categorical), as they did for the checklists (check=yes, no check=no), you end up having to consider only one value, e.g. all *yes* answers, and you compare raw values or percentages. If you use inferential statistics, logistic regression is normally used here.

Comments exercise 3

Lexical item is an important factor here as /ɑ/ and /ɔ/ are not always merged. In Canada, *sorry* does not normally sound like *sari* /ɑ/. Self-reporting seems highly reliable for the low-back merger (note, however, that for yod dropping, results were reliable for some words, e.g. *avenue*, but less so for others, e.g. *student*). There is a match between written questionnaire and the acoustic analysis in seven out of ten (*cot/caught*) and nine out of ten cases (*Don/Dawn*, and *sorry/sari*). Nonetheless, it is clear that for some words reliability is better than for other words. Dollinger (2012: 96) also discusses individual speaker behaviour in his analysis. In particular Anton was very nervous during the interviews, which may explain his mismatches. Alternatively, he may maintain a distinction between the sounds in production but not in perception.

CHAPTER 9

Comments exercise 1

Variable features are overgeneralised here, and particular variants are assigned to particular locales without acknowledging their local variability (only the use of the word *often* hints at variability). A statement like this cannot be taken for granted. Its value lies in its indication that certain features probably did occur at a certain point in a particular locale (although we can infer little about frequencies), and in its ideological dimension in that pronunciations such as *nothin* and *nothink* are seen as wrong and “barbarous” – rather than regional (cf. Montgomery 2012).

In Normative Linguistics, phonetic and phonological features are of course of a slightly different nature than other features as they are methodologically more difficult to deal with. While they can be perfectly well investigated in prescriptive work (e.g. see Beal 2004), aligning pronunciations with spelling in actual language use is much more difficult.

In contrast to syntactic variation, which is relatively easily recognised in historic texts, pronunciation is much harder to study. It is not always consistently replicated, and, as spelling becomes more codified (first regionally and then nationally), one particular spelling may hide a variety of different pronunciations. Compare for example the word *herbs*, which is pronounced with an [h] in RP, while in most American varieties, it isn't. Nonetheless, a researcher reading 21st century texts 200 years from now would never be able to detect this regional pattern as the spelling rarely varies. Historical texts do tend to be more variable in the representation of pronunciation than current texts, however.

Comments exercise 2

To test whether perceived dialect areas, identification and evaluation of actual New Zealand accents overlap, you could first conduct a rating task of dialect areas based on the composite map. This could be followed up with a rating task of voices from several NZ locations: those mentioned on the map. Accents are rated for correctness and pleasantness. This should also be map-based and could incorporate an identification task to ensure listeners know which accent they are hearing.

The draw-a-map task allows us to gain access to non-linguists' perceptions of dialect areas (as a composite perceptual map) and how they are evaluated. This map can now be related to ratings of voice samples from selected areas and similarities and differences can be established. The identification task brings in yet another dimension of comparison. It allows us to uncover whether listeners can also identify dialects from voice recordings and to what extent evaluation of voices, identification of voices and perceptual dialect maps overlap.

References

Beal, Joan. 2004. Marks of disgrace: Attitudes to non-standard pronunciation in 18th-century English pronouncing dictionaries. In Roger Lass and Marina Dossena (eds.) *Methods and Data in English Historical Dialectology*. Bern:

Peter Lang, 329-349.

Montgomery, Chris. 2012. Perceptions of dialects: Changing attitudes and ideologies. In Terttu Nevalainen and Elizabeth Closs Traugott (eds.) *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 457–469.

CHAPTER 10

Comments exercise 1

Many different outcomes are possible here. We can only give a few examples.

There are at least five tokens of word-initial respellings where a <d> is used in place of a standard English <th>. This was not done in “the hevil”.

There are at least four tokens of word-initial respellings where a <t> is used in place of a standard English <th>. This was not done in *both*.

There is at least one token (*cold*) where we would normally find a word-final <d> in standard English, which is represented as <'> here.

There are at least two tokens (*hevil*, *herased*) of word-initial respellings where an <h> is written before <e>, where standard English would not normally have an <h>. There are many exceptions.

Smith is not consistent, and she does not only mark features particular to patois. For example, t/d deletion, as in *cold*, and th-stopping, as in *the*, occur in many varieties of English.

Questions 3, 5, 6-8 ask for students opinions, so we do not provide comments for these.

Comments exercise 2

The conventions are roughly similar but there are some differences. For example, the use of square brackets and how pauses are noted differs. Confusion may potentially arise if a reader read the square brackets in an unintended way.

Another important difference lies in how intonation has been marked. Holmes and Schnurr (2006) only use the question mark, and it represents rising or question intonation. Any other intonation patterns are not explicitly marked but are defined by absence of a rise. Also note that a question mark does not mean the exact same thing in Holmes and Schnurr (2006) and in Bucholtz (1999). A question with falling intonation could be marked with a question mark in Holmes and Schnurr but not in Bucholtz.

CHAPTER 11

Comments exercise

(a) Your table should be set up something like this:

Example	Verb	Quote type	Quote content	Variety of Eng
I'm like "uuups"	like	mimetic	ambiguous	UK
She goes "well Frank opened his big mouth"	go	linguistic	speech	USA
And he goes "do you want to dance?"	go	linguistic	speech	USA
I go "no, no"	go	linguistic	speech	USA
∅ "of course you can"	zero	linguistic	speech	USA
I mean I was like trapped, rather like being a rabbit in the headlight you know, it was like "ahhhh"	like	mimetic	thought	UK

Note that there are twenty examples of verbs of quotation in the 16 examples. Note also that some of the examples are ambiguous or can't be classified without the larger context that they occurred in. Buchstaller would have had access to this, but we don't, so we may need a 'don't know' or 'ambiguous' code for some variables.

(b) You might have two summary tables, one for each independent variable (we won't fill them out with numbers because that's half the point of the task, but this will get you started if you're stuck). They might look like this:

Verb of quotation	Content of the quote			Total
	Reported speech	Reported thought	Ambiguous / unclear	
(be) like				
go				
zero				
say				
Total				

Verb of quotation	Quote type		Total
	Mimetic	Linguistic	
(be) like			
Go			
Zero			
Say			
Total			

(c) Looking at the results for these 20 tokens of verbs of quotation, you might be struck by other things that you could code for, e.g. the subject of the verb (*I, he, my daughter, it, etc.*) or the tense of the verb (past, non-past).

Buchstaller has tried to give examples of both *go* and *(be) like* being used across both the independent variables and both varieties of English, but in a more natural sample, you might start to see patterns that suggest one form is associated with UK English more than USA English or vice versa. You might also start to see patterns that suggest *(be) like* is associated more with reported thought than other verbs are.

CHAPTER 12

Comments exercise 1

Labov 1972: The dependent variable is the presence or absence of [r], and the independent variables, as Labov lists them, are “the store, floor within the store, sex, age (estimated in units of five years), occupation (floorwalker, sales, cashier, stockboy), race”, and “foreign or regional accent, if any” (173). All variables are categorical. Note that age can be continuous, but Labov has made it categorical.

Eckert 1988: The dependent variable in this paper is the pronunciation of the vowel /ʌ/. Eckert coded each vowel token in her data as conservative, extremely innovative, or intermediate, thus taking a measure that could be continuous (formant frequency) and making it categorical. She examines the independent variables of parents’ socioeconomic class (Working Class, Lower Middle Class, Upper Middle Class), social category affiliation (Jock, Burnout, In-between), and location (Redford, Livonia).

Labov 1990: This section of the paper examines two continuous dependent variables, the first and second formants for a number of vowels. Independent variables are (not surprisingly given the title of the article) sex and social class (unskilled, skilled, clerical, managerial, professional), as well as age and the phonological environment of many of the vowels (e.g. in a closed vs. open syllable, followed by a nasal). Each of these independent variables, apart from age, is categorical.

Comments exercise 2

Some of the patterns we found in the data are:

Women in this data use *be like* more often than other quotatives (51% vs. 49%), while men use *be like* less than other quotatives (47% vs. 53%). However, a chi-square test does not find this difference between the sexes to be statistically significant ($\chi^2(1, N = 500) = 0.93, p = 0.34$).

Like Tagliamonte & D’Arcy (2004), we found that *be like* was used more often with first person subjects (59%) than with third person subjects (45%). This difference is statistically significant ($\chi^2(1, N = 451) = 8.48, p = 0.004$). We carried out this analysis by opposing *be like* to all other quotatives, grouping 1.sg tokens with 1.pl tokens and 3.sg tokens with 3.pl tokens, and omitting data from all other subject types, but you could also consider grouping noun phrase subjects in with 3rd person ones. In that case, the rate of *be like* use with 3rd person subjects becomes 43%; this is still a significant difference from *be like* use with 1st person subjects ($\chi^2(1, N = 479) = 10.71, p = 0.001$).

Also like Tagliamonte & D’Arcy (2004), we found that *be like* was used more with internal dialogue (83%) than with direct speech (40%), another statistically significant difference ($\chi^2(1, N = 495) = 59.87, p = 1.01 \times 10^{-14}$). For this analysis, we opposed *be like* to all other verbs of quotation, grouped *said aloud* and *imitation* tokens together to form a *direct speech* category, and grouped *thought* and *non-verbal feeling* tokens together to form an *internal dialogue* category.

One weakness you could see in this data is that noun phrase subjects have not been coded for whether the noun phrase was singular or plural, in the way that pronoun subjects have. We suspect that subject plurality does not play a strong role in quotative variation in this data set: in pronouns, at least, singular vs. plural do not significantly differ in *be like* rate, with 50% *be like* use in singulars and 56% in plurals ($\chi^2(1, N = 467) = 0.29, p = 0.59$). Still, this is a nice reminder of how important it is to code your data in detail for as many possible variables as you can (ch. 11): had we coded noun phrase subjects for number from the outset, we could look at this easily, but as it stands, there's some painful recoding in our future should we wish to examine this.

References

Tagliamonte, Sali, and Alexandra D'Arcy. 2004. *He's like, she's like*: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8: 493–514.

CHAPTER 13

Comments exercise 1

Graph 1

This graph contains redundant percentages. (ING) is a binary variable, with two variants, [ɪn] and [ɪŋ]. If a group uses 99% [ɪn], they necessarily use 1% [ɪŋ]. There is thus no reason to graph them both.

The y-axis on this graph goes up to 120%, an impossible value. Excel will do this by default if one of your percentages is particularly close to 100. This should be corrected so that the y-axis stops at 100%.

The abbreviations 'LWC' and 'UWC' probably stand for 'Lower Working Class' and 'Upper Working Class', but there's no reason to make your reader guess this. Spell out abbreviations unless you provide a key to them in the caption, or have made it very clear in your text what they stand for.

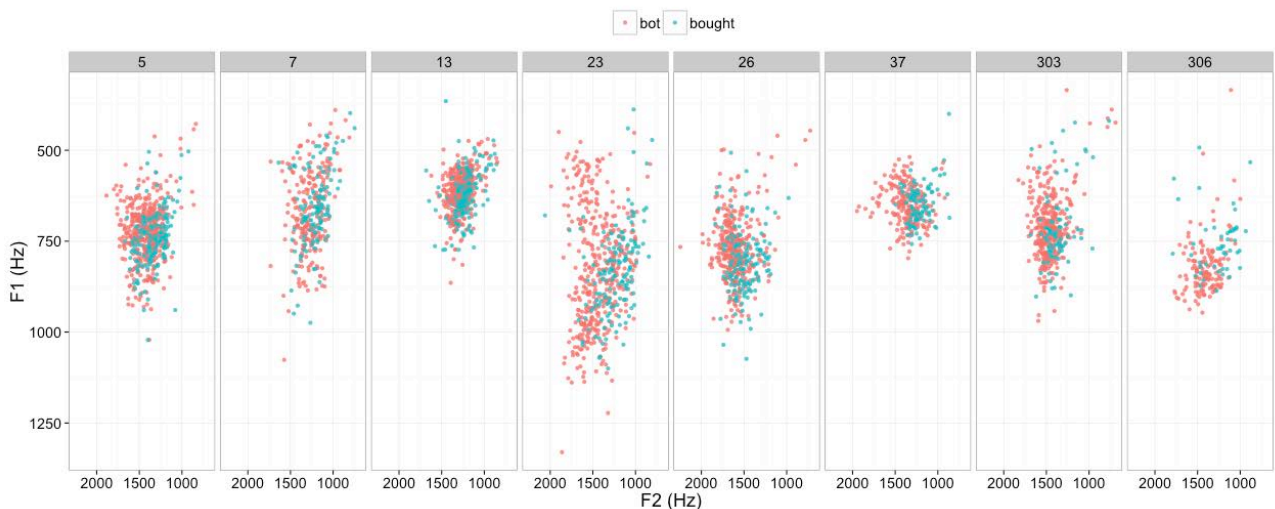
Graph 2

Look at all those bars! There's no reason to make this a bar graph rather than a line graph. The x-axis is numeric, and clearly ordered from lowest to highest. Additionally, there must be fractional values of the variable represented on the x-axis in between each tick for which we could interpolate the values of the variable represented on the y-axis.

Due to a lack of labeling, it's completely unclear what this graph is showing. What's on the x-axis? It's something numeric, and age is a good guess, but it shouldn't be your reader's responsibility to figure that out. And what's on the y-axis? Clearly it's the percentage of something, but percentage of what? The legend, which just says 'Series 1', a label that Excel provides by default, is completely uninformative. When you plot one variant of a binary variable, you must make completely clear from the graph which variant you are plotting.

Comments exercise 2

The most straightforward way to visualise vowels is with an F1/F2 scatterplot. Below, we've created one scatter plot per speaker. (Speakers are identified with a number.)



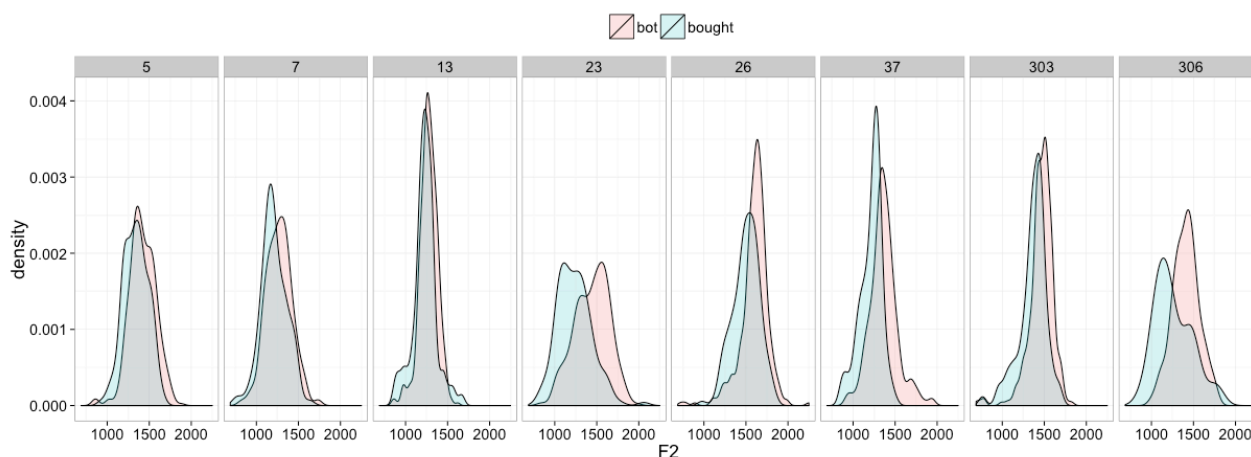
Based on these scatterplots, no speaker appears to have a complete separation of *bot* and *bought* in production: they all show a considerable amount of overlap in the two vowels. Some show less overlap than others, though: compare, for instance, the distribution of points for Speaker 23, who seems to have some separation of the two vowels in the F2 dimension, to that of Speaker 26, for whom any separation is much less obvious.

Other means of visualising the data are necessary when you want to focus on only one parameter, say, F2. Compared to simply eyeballing a scatter plot, histograms can give us a better idea of whether one vowel category has more F2 measurements in a particular range of values than the other vowel category does. The histograms below represent the counts of the various F2 values for each speaker.

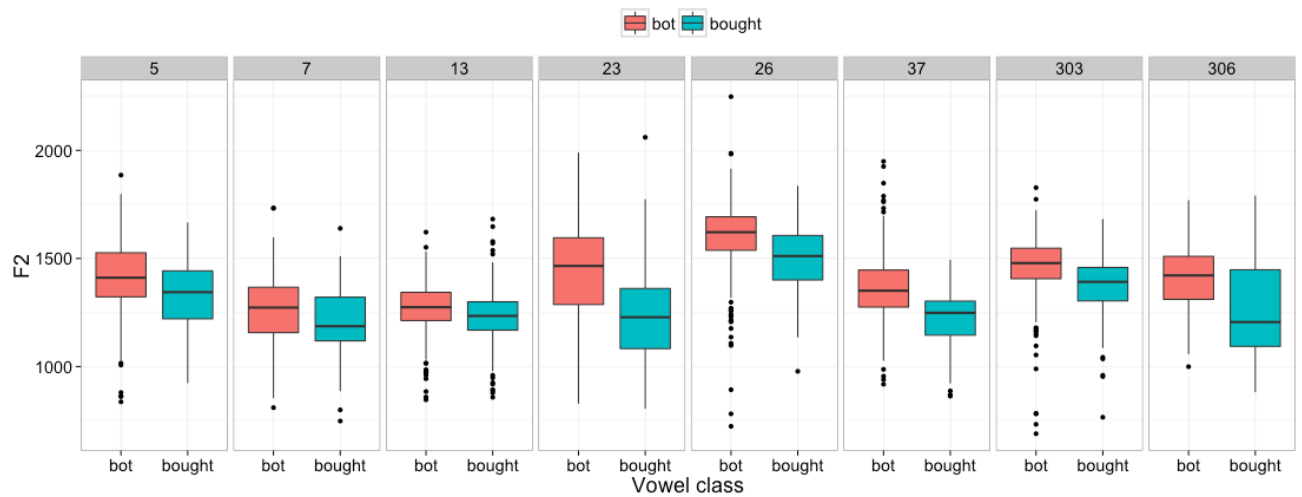


From the histograms, we can see that several speakers' F2 distributions are very similar for the two vowels, but for some speakers, like Speaker 23, there are more *bought* tokens in the lower ranges of F2 than *bot* tokens.

Density plots are similar to histograms, but rather than grouping the observations into bins of arbitrary size, they plot the probability density: the probability per unit value of the variable being plotted (in this case, Hz) of getting a particular value. The density plots below allow us to see even more clearly the difference in distributions between, say, Speaker 13 and Speaker 23.



Histograms and density plots allow easy comparisons of distributions across speakers or categories. Boxplots, like the ones below, make it somewhat harder to compare distributions just by eyeballing, but allow quick and easy comparison of central tendencies (namely, medians).



Not all of these plot types are readily available in Microsoft Excel. We've made ours in the free statistics program R, using the ggplot2 add-on. Your university may have other statistical software that you can use to create plots like these.

CHAPTER 14

Comments exercise 1

The dependent variable is continuous (specifically, ‘normalized Euclidean distance measurements of FACE and GOAT tokens’). Even if you don’t know what Euclidean distance is, the use of linear (as opposed to logistic) regression and the words ‘normalized’ and ‘measurements’ clue you in to a continuous dependent variable.

Random effects are speaker and lexical root, and fixed effects are attitudinal index score, style, sex, age, vowel duration (of which they have taken the natural log), and following/preceding voicing, manner, and place of articulation. Speaker and lexical root are both categorical independent variables that couldn’t be sampled in their entirety, so they are good candidates for random effects.

The researchers find that the dependent variable is significantly affected by three interactions: the interaction between preceding sound and age group, the interaction between following sound and age group, and the interaction between log-duration and age group.

Comments exercise A

In our Rbrul analysis, we examined fixed effects of sex, age, subject, tense, and content, as well as an interaction between age and sex (to investigate past findings, e.g. Labov 1990, that a difference between the sexes may not be apparent in the earliest stages of a change, i.e. among the oldest speakers in apparent time). We found significant effects of content and age, as follows:

BEST STEP-DOWN MODEL OF RESPONSE verb.of.quotation IS WITH PREDICTOR(S): content (8.74e-13) + age (1.77e-10)
[p-values dropping from full model]

```
$content
      factor logodds tokens be like/be like+other centered factor weight
      thought  1.637    79           0.899
non-verbal feeling -0.018    20           0.650      0.495
      said aloud -0.749   277           0.401      0.321
      imitation -0.870   100           0.430      0.295

$age
continuous logodds
+1 -0.055
```

Subject did not come out to be significant, even though we found a significant effect of subject on quotative choice when we performed a chi-square test for chapter 12’s exercise. Presumably that subject effect was tied up with some other effect, which Rbrul has been able to sort out: perhaps the higher rate of use of *be like* we found with first person subjects was better attributable to the higher rate of use of *be like* with thoughts (which presumably tend to take first person subjects).

Comments exercise B

Rbrul has found significant effects of age, gender, and preceding segment. Note that Rbrul does not generate factor weights when analysing a continuous dependent variable, but we can figure out the nature of the effects of significant predictors by looking at the provided means and their associated coefficients

(headed "coef"). The segments t/d are found to be associated with a more back vowel, as shown by the lower mean and the negative coefficient (indicating a negative effect on F2: i.e., lower numbers). The same can be said for males (also showing a negative coefficient and a lower mean). The result for age can be interpreted as follows: with each one-unit increase in age (so, with each year), F2 decreases by 2.268 Hz (the value in the "coef" column). In other words, F2 decreases as speakers get older: fronting is a change in progress.

References

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205–254.

CHAPTER 15

Comments exercise 1

Here is how we marked up the text.

yellow = qualitative phrase or comment
blue = quantitative phrase or comment
green = both?

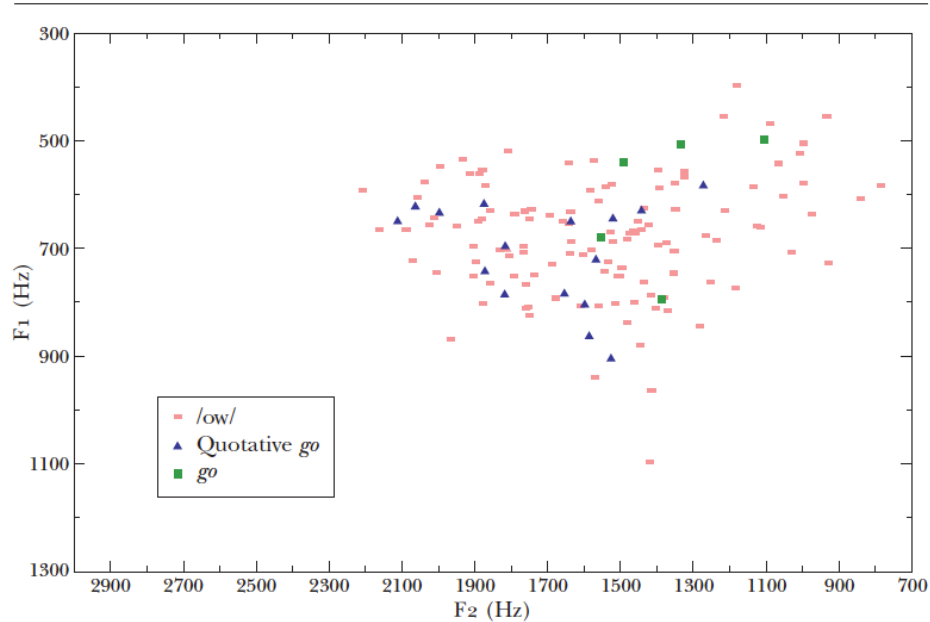
“In what follows, I will focus on the relation between Rachel’s teenage persona and the fronting of /ow/.

A clear example of the function of fronting is offered by Rachel’s pronunciation of *go*, which serves as a common quotative in Rachel’s speech. Use of quotation in narrative is a well-known way of showing involvement—of making narratives come to life and creating voices rather than simply reciting actions (e.g., Schiffrin 1981; Tannen 1982). Quotation takes on increased importance in preadolescence as narrative becomes a central genre in social drama and an important means of enacting a teenage, crowd-member stance. There are 16 quotative occurrences of *go* in the speech sample under study here, and all of them introduce narratives about Rachel and her peers, more specifically about their relationships and events in the crowd. One can, therefore, take the use of quotative *go* as indexing aspects of her crowd status. For example, in the following, Rachel is telling me why she doesn’t like Brenda, a girl in the crowd. Apparently Brenda has written something dirty in the slam book that is going around the crowd. A slam book is a notebook initiated by one or a few individuals. It contains one or more questions and circulates among peers for them to write their answers. This slam book had a bunch of questions about favorite things, one of which was “favorite body part.” Harry, a popular boy, told Rachel that Brenda had tried to convince him to write dirty stuff:

And she’s all, “No, put grosser, like, really disgusting body parts.” Um. And he’s all, uh, “I don’t know, I don’t wanna say.” And then, she’s all, “How about dick.” And then, he’s all, “Whatever. I don’t, you know, I don’t wanna get into this.” [. . .] But nobody knows that. You know, Harry told me. But I would believe Harry, but I would never believe Brenda. Brenda lies. Like Brenda, when I was playing, um, tag, she goes, um, cuz she got tagged, and then she–, um cuz she was in there for four times, and um, and then she goes, but Kristin tags her, and then she goes back to me and she goes, “Rachel, that wasn’t a fair tag cuz you just went like this to me. And, you tagged me bare–, barely. And, I, you hardly even tagged me.” I go, “Excuse me, Kristin tagged you.” She’s like, “Oh. Kristin, you, you tagged, you tagged me the wrong way.” I go, “What is she doing? She’s like totally lying.”

Figure 1 shows all 142 measurable occurrences of /ow/ in this interview. As the figure shows, Rachel’s occurrences of /ow/ in quotative *go* are significantly more fronted than those in nonquotative occurrences of *go* (F2 $t[20] = -3.686, p < .001$; F1 $t[20] = -2.479, p < .01$).

FIGURE 1
Rachel's Fronting of the Nucleus of /ow/



In fact, occurrences of quotative *go* are significantly more fronted than all other occurrences of /ow/, including those following apicals ($F2\ t[141] = 2.114, p < .025$.)" (Eckert 2011: 93-95)

Comments exercise 2

(a) Meyerhoff starts out with qualitative observations about who did and didn't use *sore* to express empathetic concern when she was doing her fieldwork in Vanuatu. She makes a qualitative suggestion, namely, that men's avoidance of it might be because the men in her sample were, in general, less oriented to others' feelings and experiences than the women were (if you read the article, even more qualitative data is used to try and say why *this* might be). Then it tries to provide a quantitative perspective on the proposed notion of 'other-orientedness' and counts the relative frequency of pronouns that refer to the addressee in women's and men's speech.

(b) Improve it? You've finished the book. You should be able to tear this one apart with glee. Go ahead. Have some fun. We did. Thanks for sticking with us. In the final chapter, we'll explain how you can write it all up beautifully.

CHAPTER 16

Comments exercise 1

Move 1 begins with “most recent” and establishes the general field very nicely and succinctly by reviewing previous research (step 3). Move 2 starts with “the present article”. Note how a research gap (step 1B) is created by comparison. This is signalled with “instead”, and it is suggested that since “the organisation of discourse” is not covered by the two research strands outlined in move 1, this represents the research gap. Move 3 starts with “following Auer”, and the phrase “we argue” signals clearly that the authors are about to present the main point of their paper. Alternatively, one could argue that filling the research gap already starts in move 2. The authors are, thus, moving from announcing the present research (step 1B) to announcing principal findings (step 2). Note how the tense changes from the discussion of previous research to the present study.

Comments exercise 2

The main inconsistencies are the following:

- (1) The list is not alphabetical: Milroy precedes Labov.
- (2) The publication date for Roberts should be in brackets or all other references should be changed to be consistent with Roberts.
- (3) In Schleeff, Meyerhoff and Clark, Clark’s surname should precede her first name, as in the other multi-author references.
- (4) Page numbers should be listed for Tagliamonte.
- (5) Only the first word in the Trudgill title should have an initial capital.
- (6) Labov’s first name should be given in full.
- (7) In the Schleeff (2013) reference *effects* should not be capitalised because words have not been capitalised in the other words following a colon.