7

Closer Looks at Speech Rate and Pause Variation: Interlocutors and Accommodation

7.1 Introduction

Chapter 5 focused on large-scale analyses of the speech rate and pause data derived from the SLAAP archive and Chapter 6 reexamined some of these data from additional perspectives to gain deeper understandings of their patterns. In order to assess overarching patterns in the data in those chapters, I downplayed potential qualitative differences between the interviews that were examined. There are also, of course, other kinds of potential influences on speakers' timing features that we have not yet considered. I have not, for instance, examined the role that the interlocutors involved in the interviews may play in influencing talkers' speech rates and pause patterns. Different kinds of interactions can occur in sociolinguistic interviews, which may yield different styles of speech (in terms of, for example, Labov's 1972 attention to speech, or Tannen's 1984[2005] conversational styles), and I have not considered issues of comparability between the interview excerpts used for analysis here (see Gregersen, Beck Nielsen, and Thøgersen 2009 and Gregersen and Barner-Rasmussen 2011 for some recent considerations of issues of comparability across sociolinguistic interviews). There are also likely important differences across the interviews that are beyond the capabilities of the present corpus-based analysis. For example, the Washington, DC data come from sociological interviews as opposed to sociolinguistic interviews (Mallinson and Kendall 2009) and, while it is possible that the difference in interview types impacts the speakers' pause and speech rates, I cannot separate this fact from the regional factor (since these interviews comprise the entire data available for Washington, DC). Based on the earlier research (reviewed in Chapter 2), it is also possible that other differences – like how well the interviewees know their

interviewers (cf. Rickford and McNair-Knox 1994) – impact aspects of speech timing in the interviews. I have proposed earlier that some differences such as the age effect for speech rate might relate to factors like shyness on the part of the youngest speakers and/or less comfort with the recording situation (and gadgetry) for the oldest speakers, and these kinds of factors too may influence the data. But the information necessary to investigate some of those potential issues is not available in the diversely collected interviews in SLAAP and must remain beyond the capabilities of the present kind of corpus-based sociophonetic analysis.

Yet, many interactional questions can be pursued via the data available here and I pursue several of these further in this chapter. I first consider, in §7.2, the influence of interlocutor attributes on the speech rates and pauses of many of the speakers from Chapter 5. Then, in §7.3, I examine a finer level of accommodation in speech timing by looking at two datasets where, in each, we have data from a single speaker in several interviews with different interlocutors. The data examined here continue to come from SLAAP and for the most part overlap with the data used in the previous chapter, although I also draw on some transcribed recordings from SLAAP that were not used in the main, corpus-based analysis. These datasets are explained further in their relevant sections.

7.2 Interlocutor effects on speech rate and pause

In this section, we take advantage of the interview nature of the recordings in SLAAP to consider potential interlocutor and accommodation effects in the speech timing data. Are there effects on these interviewees' pauses and speech rates based on the properties of their *interviewers*? Continuing with a macro-level focus, we ask whether the sex and ethnicity of the interviewers matter, and further, whether the number of participants in an interview influence the speech rates and pause durations of the interviewees. As was introduced in Chapter 2, much of the social psychological research on rate of speech has indicated that speech timing can be impacted by properties of one's interlocutors (Jaffe and Feldstein 1970, Feldstein, Alberta, and BenDebba 1979, Siegman 1979a) and that it is a likely site for convergence across speakers in interaction (e.g. Giles, Coupland, and Coupland 1991; see also Pardo 2006 on phonetic convergence). Much of the accommodation literature has focused on the meaningfulness of accommodation – the role that it plays in managing interpersonal interactions and so forth – but it should be noted that some studies have indicated that speech rate accommodation might be even more pervasive than these social psycholinguistic theories predict. Staum Casasanto, Jasmin, and Casasanto (2010), for instance, found that speakers accommodate in their speech rate to that of a "virtual interlocutor," a computer avatar, despite knowing that the avatar is not human. With evidence in the literature like this, we have strong reason to anticipate finding accommodation effects in the data available here.

We begin examining these questions through the speaker-level median values rather than the utterance-level data. As we have seen, the median values provide a useful dataset for statistical analysis, and considering the influence of interaction-related factors at a speaker level seems appropriate here. The examinations in this section deal with a subset of the data examined in Chapter 5. These are the overall median values for 146 of the speakers – all of the *interviewees* examined in Chapter 5 with the exception of four (for whom the data came from changing sets of interviewers). In other words, we now examine 146 interviewees from sociolinguistic recordings and we do so with an eye to their interviewers and coparticipants. To further describe the data used here, and to confirm that this subsetting of the data does not impact the comparability with the analyses presented earlier, Table 7.1 presents the means of these 146 subjects and compares them to the means of the original 159 subjects by ethnicity (ordered alphabetically). Note that none of the differences are near significant.

The sociolinguistic interviews in SLAAP, which comprise the main data of these studies, have different configurations of participants. The majority of the interviews examined here are of the cardinal interview type, having a single interviewee and a single interviewer. Many of the interviews, however, are not just dyads; sometimes there are two interviewers with a single interviewee and sometimes there are two

Table 7.1 Minor and nonsignificant differences between subset and main data

	Subset N (diff.)	Subset mean speech rate (σ/sec)	Overall mean speech rate (σ/sec)	Speech rate t-test results	Subset mean pause duration (ms)	Overall mean pause duration (ms)	Pause duration t-test results
Afr. Am.	51 (–4)	4.64	4.63	$p = 0.94$	408	409	$p = 0.93$
Euro. Am.	43 (–5)	5.04	5.07	$p = 0.77$	406	404	$p = 0.93$
Latino	40 (–2)	4.79	4.74	$p = 0.77$	424	422	$p = 0.93$
Lumbee	12 (–2)	4.97	4.99	$p = 0.94$	414	422	$p = 0.87$

interviewees interviewed by a single fieldworker, and so on. We can hypothesize that this factor might matter and that participants in the interviews may exhibit speech timing features that are, on average, influenced by the number of people with whom they are speaking. We begin then by examining the effect of the number of participants in an interaction on an interviewee's (median) speech rate and pause duration. I use the term "participant" loosely here to indicate someone in the interaction irrespective of whether she or he is an interviewer, an interviewee, or even an interloper who has joined the conversation uninvited.[1]

Figure 7.1 displays boxplots for speech rate and pause organized by the number of participants. Through these boxplots, we see a pattern whereby speech rate decreases as the number of participants increases from two to four and then sharply increases for interviewees in interactions with five participants. This pattern is slightly evident for pause duration as well, but as in the previous two chapters, we see less patterning in the pause data.

An ANOVA finds the participant number pattern significant for articulation rate ($F(3, 142) = 3.31$; $p = 0.022$). A post-hoc Tukey test indicates that at closer inspection only the comparison between five and four participants is actually significant ($p = 0.040$). The comparison between four and two participants just fails to reach significance ($p = 0.060$) but the rest of the comparisons are far from significant. The pattern seen here is interesting and interpretable. For low numbers of

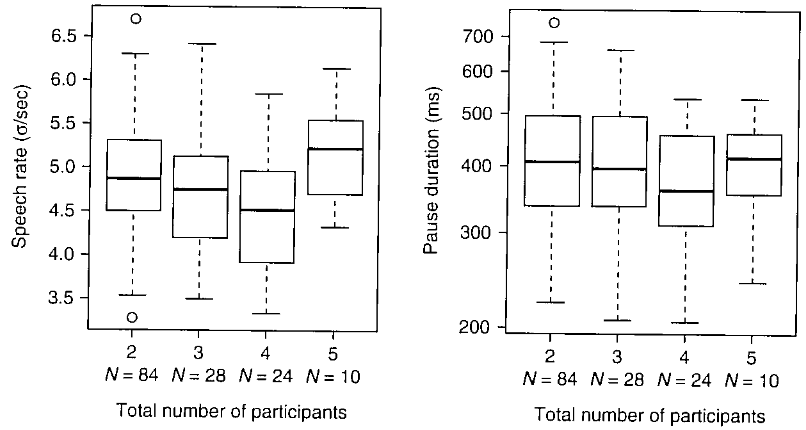

Figure 7.1 Effect of number of participants on articulation rate and pause duration

total participants, speakers' speech rates decrease as more interlocutors enter the conversation. This can be seen, perhaps, as a sort of deferentiality or a strategy for ensuring that all hearers can follow one's talk and all speakers have opportunities to share the floor. At a certain point (as we see here at five participants), however, the increased competition for talk time created by the higher number of interlocutors may cause speakers to speak faster, to fit more contribution into their more limited talk time. That is, with a large enough number of participants, talk may become a competitive enterprise more than a collaborative enterprise. This is of course speculative, but the pattern begs further consideration in the future. As expected from the pause boxplots, ANOVAs for the pause data (both raw and logged duration) do not obtain significance. Given the striking articulation rate pattern (however we potentially interpret it), it is notable that the pause durations do not follow suit. We do not see evidence of pause patterns shifting to accommodate the different numbers of participants.

In order to investigate whether the sex of speakers' interlocutors might influence their speech rate and pause, I coded each interviewee for the sex of their interviewer(s) and/or coparticipant(s). These were marked as "female" if all the interviewers and coparticipants were female, "male" if all the interviewers and coparticipants were male, or "mixed" if the interviewers and coparticipants comprised males *and* females – that is, if (other than the participant) there were both males and females taking part in the interaction. "Mixed" interactions necessarily had at least two interviewers or coparticipants in addition to the participant. (Interactions marked as "male" or "female" had one to four participants in addition to the interviewee.) I also coded as a separate potential factor whether the sexes of the interviewees were the "same" or "different" than the sexes of their coparticipants. Where there were both male and female (i.e. "mixed") interviewers, this value was coded as "different."

For these sex data, the interesting patterns arise from a full consideration of the actual sexes of the participants (more than the simpler "same" vs "different" categorization). Figure 7.2 provides boxplots for the speech rate data, with each boxplot providing data for a pairing of the interviewers' sex (labeled as Intvers in the figure) and the interviewees' sex (IntvEEs). We see in this figure that speech rates are lower for interviewees who were interviewed by only females.

A *t*-test finds the difference between the subjects interviewed by only females (mean = 4.46 σ/sec) and the remaining subjects (mean = 5.05 σ/sec) to be highly significant ($p < 0.000001$). The difference here, of 12 percent, is well above the 5 percent JND threshold, indicating that these

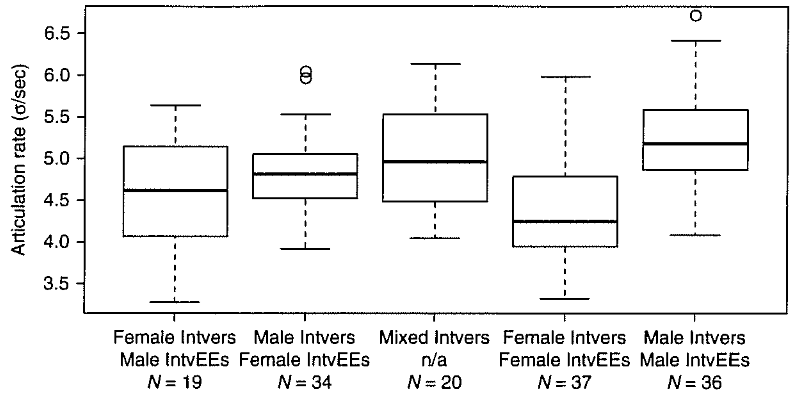

Figure 7.2 Effect of interviewer and interviewee sex on articulation rate

differences should be quite perceivable to hearers (see §2.2 and §5.6; Quené 2007). At the same time, the difference between the speech rates of the female subjects who were interviewed by females and the male subjects interviewed by females is not significant ($p = 0.385$), indicating that the important factor here is primarily the sex of the interviewers, not the relationship between the interviewers' and interviewees' sexes.

Figure 7.3 provides similar boxplots for the pause duration data. As with the data for number of participants, the pattern found for speech rate does not appear evident here. The only visible difference among these data is that male subjects interviewed by females have longer pauses (mean = 476 ms) than the other groups (mean = 403 ms). This difference is confirmed by a *t*-test (conducted on log-transformed durations; $p = 0.01$) and is striking (especially as it is the only such significant accommodation effect for the pause duration data). Without other interlocutor effects on pause durations, it is hard to interpret this difference and the pattern will be left for future consideration.

We now turn to consider the role of interviewer ethnicity on the interviewees' data. The data were treated in a similar fashion as the interviewer sex factors, with each data point coded for two interviewer-related ethnicity factors. First I coded the interviewer's ethnicity (either "European American," "African American," or "Lumbee" – there were no Latino/a interviewers and in cases where there were more than one interviewer the interviewers all shared the same ethnicity). Second, I coded the data for whether the interviewee was the same ethnicity as the interviewer, as for sex, "same" or "different." Due to the larger spread

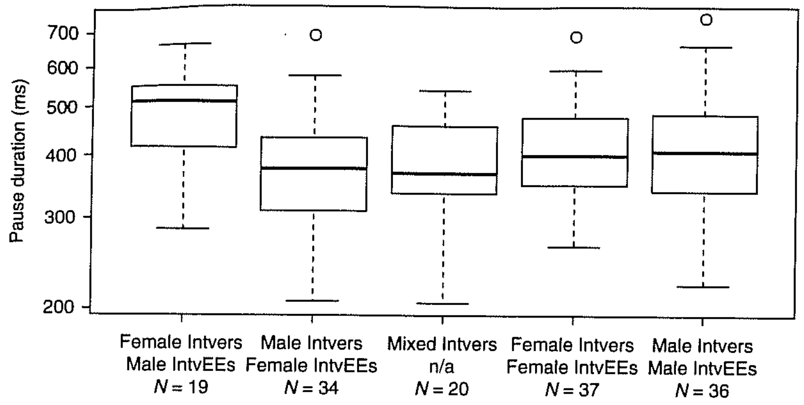

Figure 7.3 Effect of interviewer and interviewee sex on pause duration

of possible comparisons between interviewee and interviewer ethnicity than was possible for sex and the number of holes in the possible comparisons (again, there were no Latino interviewers and, in fact, only two interviews conducted by Lumbees), I present these data using solely the same ethnicity vs different ethnicity factor. This is shown in Figure 7.4. For speech rate, we notice that speakers appear to speak more slowly when interviewed by someone of a different ethnicity (mean = 4.72 σ/sec vs mean = 4.98 σ/sec), an observation confirmed by a *t*-test (p = 0.020). Pause duration, on the other hand, is not observably different and this is also confirmed by a *t*-test.

The observations made thus far, despite using some basic statistical methods to confirm the visible patterns, do not account for the interplay of all of the available factors. To test the influence of these three interlocutor-based factors – number of participants, interviewer sex, and interviewer ethnicity – along with the factors found to be significant earlier (in §5.4), I built fixed-effect linear models on these 146 speakers' speech rate and pause medians. Modeling began with the best models from §5.4 (Table 5.5, for speech rate, and Table 5.6, for pause duration) and then tested adding each of these interlocutor factors. None of the factors, however, were found to significantly improve the model fit over the earlier fixed-effect models. For pause duration, surprisingly, the same versus different ethnicity interviewer factor almost reached significance (p = 0.058), with some indication that pauses were slightly *longer* when the interviewer and interviewee shared the same ethnicity when the other factors were considered. This, while not quite significant, also

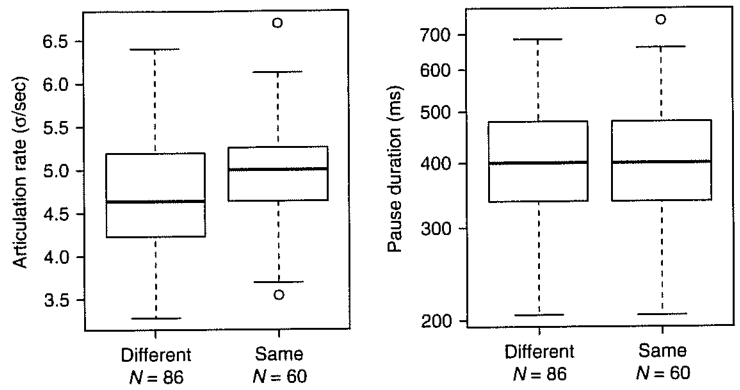

Figure 7.4 Effects of different/same ethnicity of interviewers and interviewees on articulation rate (left) and pause duration (right)

runs counter to an expectation based on assumed directions of accommodation effects and, once again, adds to the evidence that the pause duration data are not very systematically patterned by social factors.

For speech rate, the statistical models found some support for the sex of the interviewers influencing the interviewees' rates but the models which included this term underperformed the models without it. Put differently, some statistical models indicated that the effect was significant, but the best models for the speaker median rate data did not include the factor. This is quite possibly the result of the full models becoming too bloated with parameters for the amount of available data (i.e. not having enough tokens to support the high number of parameters). In order to investigate this further, I tested mixed-effect regressions on the utterance-level articulation rate data for these speakers to see if any of these participant-related factors improved the utterance-level fits. As usual, I began with the best model that was fit to the full utterance-level data in §5.3.1 and Table 5.2 and tested the addition of each of these new factors using likelihood ratio tests. This found a significant improvement in the model by adding the sex of the interviewer, INTVRSEX. The final, best model, after trimming outliers, is presented in Table 7.2 (501 measurements, 1.8 percent of the data, were removed by trimming, leaving 26,975 utterance measurements in the data). Interviewees with male interviewers are found to have significantly faster predicted rates than those with female interviewers (estimate = 0.20 σ/sec, p = 0.020). Interviewees with mixed sex interviewers also have faster estimates than those with female interviewers alone, but here the comparison is

Table 7.2 Best mixed-effect model for (trimmed) utterance-level articulation rates after interlocutor factors added

Factor	Estimate	Std err.	*p*
(Intercept)	–0.218	0.346	–
NumSyls	1.046	0.012	0.0001
NumSyls'	–5.695	0.123	0.0001
NumSyls''	15.018	0.439	0.0001
NumSyls'''	–10.096	0.430	0.0001
Age	0.046	0.021	0.0168
Age'	–1.968	0.922	0.0228
Age''	2.129	1.001	0.0232
Region = Eastern NC	0.340	0.125	0.0032
Region = Ohio	0.300	0.153	0.0306
Region = Southern NC	0.267	0.125	0.0208
Region = Texas	0.222	0.118	0.0410
Region = Washington, DC	0.261	0.160	[0.0752]
Region = Western NC	0.156	0.141	[0.2132]
Sex = male	0.282	0.109	0.0050
IntvrSex = male	0.200	0.094	0.0196
IntvrSex = mixed	0.137	0.113	[0.1936]
Ethnicity = Latino/a	0.286	0.148	0.0370
Ethnicity = Lumbee	–0.142	0.212	[0.4722]
Ethnicity = European Am.	0.337	0.120	0.0032
Ethnicity = Latino/a × Sex = male	0.056	0.160	[0.6946]
Ethnicity = Lumbee × Sex = male	0.093	0.245	[0.6672]
Ethnicity = Eur. Am. × Sex = male	–0.418	0.160	0.0064

$R^2 = 0.644$.

not significant ($p = 0.19$).[2] Overall, we find support that the sex of the interviewers impacts the speech rates of the interviewees.

All in all, the findings of this section raise some important questions about the effect of interviewers and coparticipants on the speech obtained in sociolinguistic interviews (cf. Rickford and McNair-Knox 1994, Hazen 2000b, Schilling-Estes 2004). It is hard to interpret the fact that female interviewers elicit significantly slower speech from subjects from this corpus-based account alone. One explanation may be that this is an outcome of accommodation. Since females have been found to have slower articulation rates than males (see §5.3.1 and §5.4.1), both males and females may reduce their rates when talking with females as a form of convergence (at least to their expectations). This suggestion is tentative and in need of further consideration, but without developing a closer sense of the properties of the specific interviewers and, quite likely, the exact relationships between the interviewers and the interviewees,

we cannot do more than speculate. Thus, with the acknowledgment that there are more to the patterns in the articulation rate data, if not also the pause data, than can be explained by looking at the speakers alone, we close this section and move away from general, large-scale patterns to look more closely at two smaller datasets, where we can gain closer insight into influences on within-speaker variation in speech timing.

7.3 Accommodation in pauses and speech rates

In the last section, we demonstrated that the person with whom a speaker is talking, even at a very coarse-grained level – for example a female versus a male interviewer – can have an influence on her or his speech rate (and, possibly, though seen there to a much lesser degree, pauses). However, in that examination of the corpus data we were not able to look at actual changes in individuals' speech based on changes in their interlocutors. We had to infer these changes from comparisons made across subjects, since the data did not include the same speakers with different interviewers. In this section we examine these phenomena further by turning to two case studies each involving single individuals in several interactions with different interlocutors. The first, §7.3.1, examines five interviews with the same interviewee conducted by different interviewers (Cullinan 2007). The second, §7.3.2, turns our attention to the interviewer for the Washington, DC interviews (Mallinson and Kendall 2009, Kendall 2010b).

7.3.1 A case study: who is interviewing EH?

In 2007, Danica Cullinan, a member of the North Carolina Language and Life Project (NCLLP), organized five interviews with "EH," an 82-year-old African American woman living in Raleigh, North Carolina, originally from Wilson County (about an hour east by car from Raleigh; both Raleigh and Wilson County are in Central NC).[3] In these interviews, Cullinan set out to examine intra-speaker variation by changing the interviewer for each interview, while controlling for as many other factors as possible, including location, time of day, audio equipment, and "energy level of the interviewee" (Cullinan 2007: 6).

Cullinan (2007) examined intra-speaker variation in pause durations, in addition to other speech features (such as rhythm and pitch differences between the interviews). She conducted this study primarily within the framework of SLAAP, and I am grateful now to be able to draw on her work to examine the stability and differences among EH's pause durations and speech rates across these five interviews, and to

Table 7.3 Interviewer information and data summary for EH

Int. #	Intr. ID	Ethn.	Sex	Age	Length of acquaintance	Interviewer's regional background	EH pause dur.	EH speech rate	Intr. pause dur.	Intr. speech rate
1	CT	Euro. Am.	Male	30	0 days	Raleigh, NC	406 (N = 132)	3.99 (N = 150)	615 (N = 35)	5.79 (N = 77)
2	LM	Afr. Am.	Male	45	6 years	Raleigh, NC	455 (N = 162)	4.01 (N = 162)	395 (N = 31)	4.15 (N = 66)
3	CB	Afr. Am	Female	62	>15 years	NYC → Raleigh, NC (at age of 22)	368 (N = 142)	3.99 (N = 176)	494 (N = 36)	3.75 (N = 52)
4	TW	Euro. Am.	Female	29	8 years	Wilmington, NC	370 (N = 156)	3.95 (N = 204)	186 (N = 28)	4.24 (N = 59)
5	DC	Euro. Am.	Female	28	4 years	Illinois → Raleigh, NC (at age of 8)	364 (N = 219)	4.35 (N = 230)	260 (N = 24)	4.52 (N=45)

further examine the effect of interviewers on speakers' pauses and speech rates. The social and demographic characteristics of Cullinan's interviewers, along with EH and their pause and speech rate data, are presented in Table 7.3. All pause durations and speech rates in Table 7.3 are median values.

Figure 7.5 shows EH's median pause durations and speech rates against those measures for her interviewers in each of the five interviews. We see here that despite differing values among – and differing genders and ethnicities of – her interviewers, EH's speech rate and pause duration remain relatively constant. An ANOVA finds her speech rate differences significant ($F(4{,}910) = 2.80$, $p = 0.025$), but a post-hoc Tukey test shows that her speech rates are only marginally different across the interviews. Comparisons between the interview with DC and the other four interviewers yields p values in the 0.05–0.08 range, but the differences are far from significant in the comparisons among the other interviewers. There are no significant differences in her pause data.

EH appears to provide evidence at least that a speaker's speech rate and pause realizations are not always influenced by her or his interlocutors. EH's interviewers have widely varying median pause durations and she does not. This is in line with the general lack of findings of social influences and interlocutor effects on pause. At the same time, through a closer inspection of the speech rate data (e.g. looking closely at Figure 7.5), we note that EH's median articulation rates are actually quite similar to her interviewers' for all but the interview by CT, who

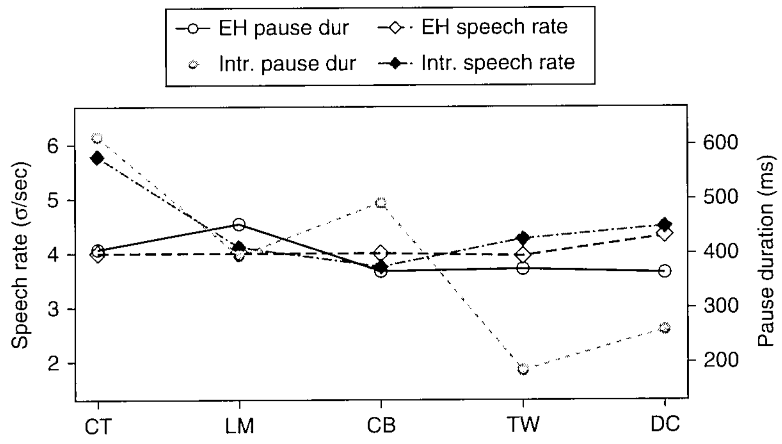

Figure 7.5 Speech rate and pause duration medians for EH and her interviewers

has an extremely fast articulation rate of 5.79 σ/sec. While we do not see evidence here of accommodation, it is possibly a result of the interviewers having too similar rates themselves (and then that CT, perhaps, is too different for convergence to take place). Thus, we can summarize from the EH data that pause durations continue to appear noisy and mostly unexplainable while speech rates are mostly inconclusive.

7.3.2 A case study: C is interviewing whom?

To examine this question further we turn to one final subset of the data. As I have mentioned in passing in earlier chapters, the recordings from Washington, DC come from an interesting source, with benefits for the project at hand. They are sociological interviews with inner-city African American adolescents conducted for a Master's project by a white woman in her mid-twenties, originally from Minnesota, who was a graduate student in sociology at North Carolina State University (Froyum Roise 2004).[4]

In 2001, the interviewer, C, worked as a counselor at a nonprofit organization called "Urban Youth Network" (a pseudonym; henceforth UYN) in Washington, DC. UYN is located in northeast DC and was founded in the 1970s to serve "at risk," troubled, and homeless youth in the city. From summer 2001 until summer 2002, C lived and worked at UYN. The next summer, in 2003, she returned to conduct an ethnographic study of the youths there. She observed 65 teenagers and interviewed 20 of them (9 boys, 11 girls). At the time the interviews were conducted, C had worked at UYN for nearly two years, and had established herself as a trustworthy adult, counselor, and confidant to the youths (Froyum Roise 2004, Mallinson and Kendall 2009). The interviews were semistructured and were designed to elicit data as to how the youth respond to the demands of inner-city life, and whether these responses differ by gender. The interviews were conducted in an office at the UYN center and were extremely similar to one another in terms of questions, scope, and setting. Since C was the sole interviewer, had a similar relationship with all of the interviewees, and conducted relatively structured – and comparable – interviews, her speech in these interviews and the speech of her interviewees provide an excellent opportunity to examine variation at the discourse level and the kinds of accommodation effects we might find for speech rate and pause variation in spoken interactions.

Twelve of these 20 interviews –10 with girls and 2 with boys – are fully transcribed in SLAAP and have been included among the 159 speakers examined in Chapter 5. As mentioned earlier, they comprise the entire data available for Washington, DC in the main analysis. Since the entire interviews have been transcribed, they also represent one of the largest coherent transcribed datasets in the SLAAP archive. Nine of these young women (all but Cal-13) and one of the young men (Edw-13) were included among the 15 speakers examined in §6.2 for having the most available data. In this section, we focus only on the interviews with the ten females, excluding the two males. We begin by examining in closer detail the data for the interviewees but then move on to examine C's speech across these ten interviews.

Table 7.4[5] presents the median speech rate and median pause duration values for each of these interviewees and Figure 7.6 displays the distributions of these speakers' speech rate and pause medians in relation to the other 149 speakers examined in Chapter 5. For our purposes we are primarily interested in the relative location and spread of the DC points relative to the distribution of the rest of the data. As we see

Table 7.4 Median pause durations and speech rates for DC females

	Median art. rate (σ/sec)	**Median pause dur. (ms)**
Ala-14	5.14 (+) N = 1,094	345 (–) N = 1,093
Asi-12	3.94 (–) N = 966	521 N = 677
Cal-13	4.19 N = 418	477 N = 299
Eli-14	4.82 N = 601	445 N = 524
Gra-13	4.74 N = 1,084	410 N = 1,173
Kei-15	5.10 (+) N = 1,123	588 (+) N = 1,077
Lat-17	4.17 N = 966	456 N = 848
Sha-14	4.38 N = 1,318	526 N = 1,142
Sha-12	3.96 (–) N = 508	505 N = 441
Shi-14	4.26 N = 1,939	416 N = 1,449
Mean	4.47 μN = 1,002	469 μN = 872
St. dev.	0.45 σN = 442	70 σN = 374

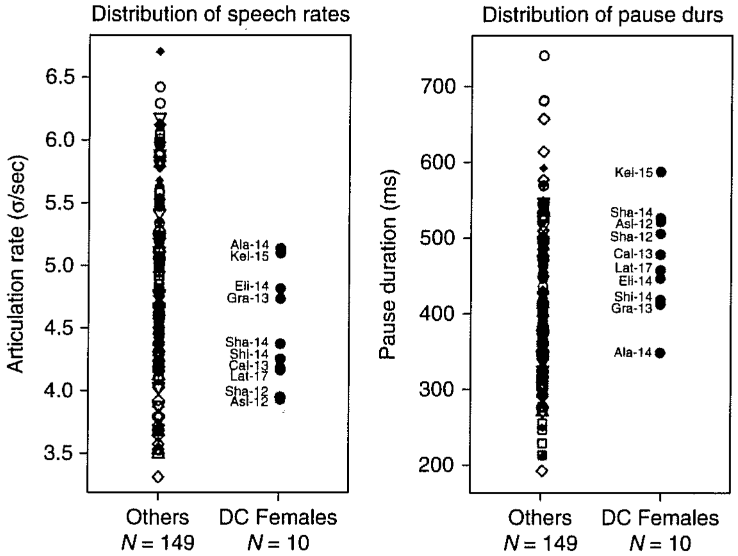

Figure 7.6 Distributions of speech rate and pause duration data for DC females

from Table 7.4 and Figure 7.6, there is a range of variation within the group, although less than we see for the other speakers. For speech rate, we see that the Washington, DC girls are at the slow end of the overall distribution. For pause, they are closer to the middle of the distribution's range, but we also note that there are actually only a few speakers in the dataset with pauses longer than the (longer) DC speakers.

As a brief aside, Figure 7.7 plots the relationship between pause and speech rate for these ten speakers. We see from this and the previous figure that there is a range of variation among the speakers' pause durations and speech rates although the range is well contained within the larger range for the rest of the speakers in the data. We also see that there is an inverse linear relationship between pause and speech rate among these speakers, provided that we ignore Kei-15, who is an outlier to this general pattern ($r = -0.79$, $p = 0.012$ without Kei-15; though for all ten speakers the correlation is not significant, with $r = -0.25$ but $p = 0.484$). This pattern is congruent with the relationship between articulation rate and pause duration (at the speaker level) seen in Figure 5.7 earlier, where we saw a slight and just significant inverse relationship between the two features (with $r = -0.16$, $p = 0.045$).

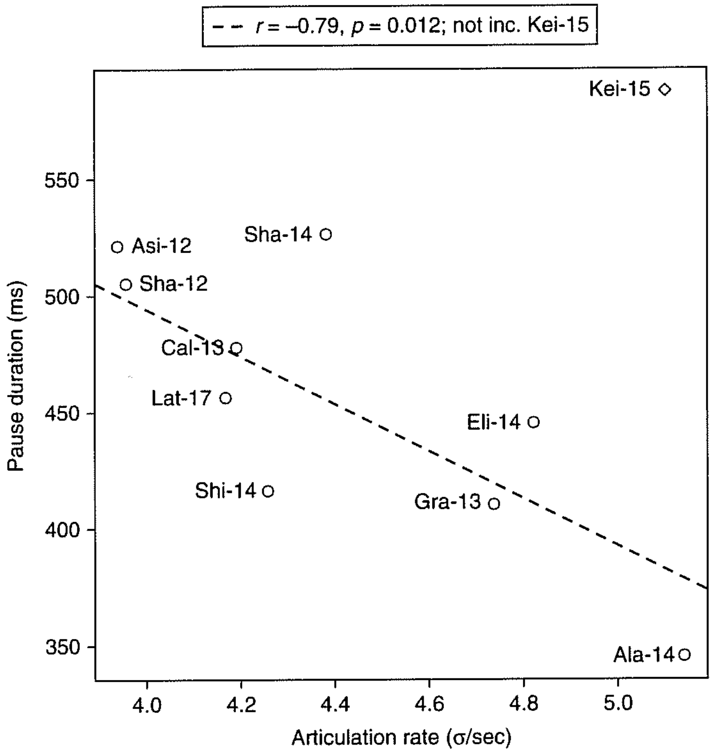

Figure 7.7 Speech rate and pause duration correlation for DC interviewees

We turn now to the main interest of this section – the patterns for C, the interviewer, across these ten interviews. Table 7.5 presents median speech rate and pause duration measures for C in each of the interviews with the ten interviewees, along with the data for those speakers from Table 7.4. Since all of the interviews were conducted in similar ways and C had similar relationships with all of the interviewees, it seems reasonable to approach these data with one of two expectations: either C's speech (i.e. her pause and speech rate) will remain relatively constant between interviews or she will show evidence of accommodation to her interviewees' speech. In comparison to the interviewees, C has a shorter overall pause duration (404 ms compared to 469 ms) and faster speech rate (4.89 σ/sec to 4.47 σ/sec).[6] *T*-tests confirm both of these differences are significant, with $p = 0.035$ for pause duration (computed based on log-transformed pause durations) and $p = 0.018$ for speech rate.

Table 7.5 Median pause duration and speech rate for DC interviewees and interviewer

	Interviewee median art. rate (σ/sec)	Interviewee median pause dur. (ms)	C median art. rate (σ/sec)	C median pause dur. (ms)
Ala-14	5.14 (+)	345 (–)	5.16 (+)	363
	N = 1,094	*N* = 1,093	*N* = 437	*N* = 391
Asi-12	3.94 (–)	521	4.72	453
	N = 966	*N* = 677	*N* = 736	*N* = 400
Cal-13	4.19	477	4.88	488 (+)
	N = 418	*N* = 299	*N* = 424	*N* = 280
Eli-14	4.82	445	5.04	383
	N = 601	*N* = 524	*N* = 432	*N* = 335
Gra-13	4.74	410	4.89	446
	N = 1,084	*N* = 1,173	*N* = 290	*N* = 199
Kei-15	5.10 (+)	588 (+)	5.10 (+)	365
	N = 1,123	*N* = 1,077	*N* = 416	*N* = 366
Lat-17	4.17	456	4.94	388
	N = 966	*N* = 848	*N* = 502	*N* = 321
Sha-14	4.38	526	4.72	387
	N = 1,318	*N* = 1,142	*N* = 637	*N* = 331
Sha-12	3.96 (–)	505	4.94	453
	N = 508	*N* = 441	*N* = 448	*N* = 346
Shi-14	4.26	416	4.51 (–)	313 (–)
	N = 1,939	*N* = 1,449	*N* = 403	*N* = 212
Mean	4.47	469	4.89	404
	μ*N* = 1,002	μ*N* = 872	μ*N* = 473	μ*N* = 318
St. dev.	0.45	70	0.20	54
	σ*N* = 442	σ*N* = 374	σ*N* = 127	σ*N* = 69

Figure 7.8 displays the relationships between C's median scores with those from her interviewees. The figure paints a picture of both stability and variability. C's pause durations vary quite a bit across interviews, while her speech rates are somewhat stable but at the same time also show some movement towards, or sometimes away from, her interviewees'. It is not obvious from the figure whether C's variability is correlated with her interviewees' speech rates and/or pause realizations. In a few cases, such as for both speech rate and pause with Ala-14 and speech rate for Kei-15, C has nearly identical median values to her interviewees, but in many other cases there is no apparent relationship, as with, for instance, Sha-12.

Finally, we look at these data from a different perspective. Figure 7.9 presents the speech rate and pause duration data as shown earlier in

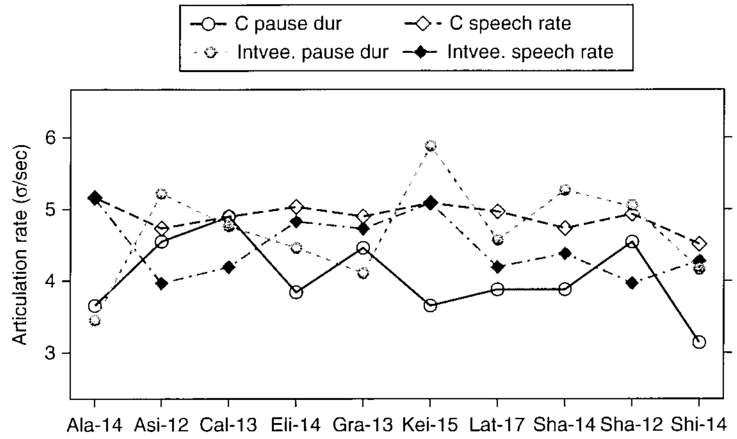

Figure 7.8 Pause duration and speech rate comparison for C and her interviewees

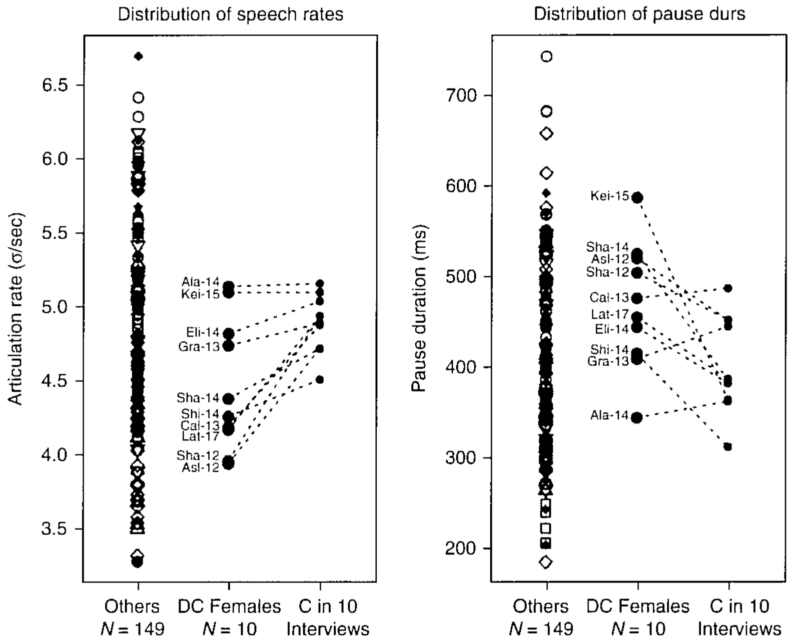

Figure 7.9 Distributions of DC speech rate and pause data, including C

Figure 7.6 for just the interviewees, but here I add the median measures for C, the interviewer, for each of the interviews as well. In the figure, the dotted lines connect C's median values with the values for her interviewee in each of the recordings. We see that C tends to speak with a higher speech rate when her interviewees do, but her overall rates are nonetheless contained within a 0.65 σ/sec range that is in the high end of the range for her interviewees (recall from Table 7.5 that her own rates have a smaller standard deviation than those for her interviewees). That is, she speaks at a faster speech rate than her slow interviewees but an equivalent rate to her fastest-talking interviewees. For speech rate there is a significant correlation in C's medians with those of her interviewees (Pearson's $r = 0.65$, $p = 0.042$). For pause, we perhaps observe a similar tendency – C often has pauses slightly shorter than her interviewees – but the actual correlation is quite low and nonsignificant due to a number of outliers (Pearson's $r = 0.22$, $p = 0.537$). While not shown, I also tested for possible correlations between C's speech rates and her interviewees' pause durations and vice versa, but these were not found to be significant.

A more complete analysis of the variation in these interviews than we have room for here, including an examination of morphosyntactic and phonological variation, would be necessary to fully understand the ways in which C accommodates to her interviewees and presents herself in each of the interviews. While this is left for now (cf. Kendall 2010b), we take from this second case study the opposite sense gained from the first (in §7.3.1). C does appear to modify her speech rate, and to a lesser degree pause realizations, in ways that relate at least somewhat to those of her interviewees.[7] At the same time, this relationship, like many we have seen, is more complex than a simple, unidimensional pattern.

7.4 Summing up

Throughout this chapter we have seen evidence that speakers' speech rates, and (as before) to a lesser extent pause durations, can vary in response to aspects specific to their interactional contexts, including the sex of their interlocutors and the number of participants in the interaction (§7.2). We found evidence of accommodation and/or self-design oriented toward specific individuals (§7.3.2), but we also found that all speakers may not be influenced by their interlocutors (or that speakers may not be influenced by their interlocutors all the time), as appeared the case for EH (§7.3.1).

How do we ameliorate the somewhat contradicting findings of the case studies of EH (§7.3.1) and C, the interviewer in the Washington, DC, interviews (§7.3.2)? EH showed little evidence of intra-speaker variation across her interviews, but C showed a patterned relationship to her interviewees for speech rate and (unaccounted for) variability in her pause durations. Despite the fact that earlier in this work (Chapters 5 and 6) I have not explicitly appreciated the difference between interviewers and interviewees – recall that 9 of the 159 speakers in the main analysis of Chapter 5 were interviewers – these two case studies highlight what is likely a major difference between these two interview roles. Interviewers have a stake in the interview process. As such, their willingness to and interest in accommodating to their interviewees is likely heightened throughout the interview in comparison to speakers with less stake in the interview. Interviewees, especially those, like EH, who participate in interviews as "subjects," often as a favor to a friend, acquaintance, or even an unknown fieldworker from a distant university, may have little engagement in the interview process, and, as such, may make little effort to accommodate to their interviewer(s) (cf. Hazen 2000b). Cullinan (2007: 6, fn. 8) describes that EH at one point reported that the interviews "wore her out." It is, perhaps, not surprising that her speech rates and pauses were little impacted by differences among her interviewers.

While this chapter found evidence of accommodation effects in the speech rate, and to a much lesser degree pause, data, the large-scale models examined in §7.2 indicated that these interlocutor-based effects are somewhat secondary and not as important as the main effects of a speaker's own, say, ethnicity, sex, and regional affiliation. Nonetheless, the discussions in this chapter, while focused on speech rate and pause, raise some questions that are important for the larger sociolinguistic picture. To what extent do these findings for pause and speech rate inform our understanding of "normal" sociolinguistic variables? Further, to what degree are pause and speech rate different from "normal" sociolinguistic variables? Are they different? There is of course so much more that could be looked at here, and this broad-brush stroke has primarily meant to strengthen, and somewhat nuance, the main effects found in the previous chapters. And, to point to areas where future work could be especially enlightening.

Part III

Speech Rate, Pause, and Sociolinguistic Variation

8
The Influence of Speech Rate and Pause on Sociolinguistic Variables

8.1 Introduction

In this final part of the book we change gears considerably from looking at speech rate and pause as the objects of study to considering their potential influences on, or roles as predictors of, the realization of other, more traditional sociolinguistic variables. Thus far, we have found more robust sociolinguistic patterns for speech rate than we have for pause. A range of factors were found to influence articulation rates and the statistical models achieved good fits (with R^2 values reaching just above 0.60). In light of all of the factors we did *not* consider (like communicative intent, discourse- and topic-related factors, affect, and so on), the modeling results were, I believe, quite impressive.

At the same time, for the pause duration data we were able to find that some social factors did influence pause durations, and that by adding some nonsocial factors to the analysis, we could improve the models. Yet, even the best models for pause were quite poor (yielding R^2 values of no higher than 0.20 and mostly below 0.10). Looking back on this apparent lack of patterning, we have to admit that for the pause data the "lawful relationship between temporal phenomena in human speech and concurrent cognitive processes" (Kowal and O'Connell 1980: 61) has largely escaped us. The kinds of cognitive processes – and likely discourse and interactional processes as well – most influencing pause patterns appear to be beyond the scope of this sort of corpus-based analysis and may require a closer attention to the individual pauses than we can give in a large-scale corpus study. In fact, I did not pursue a closer analysis of the pause data than that in §6.5 in the belief that any post hoc analysis would fail to be truly accountable to the pause data. Attempting to code the pauses in the aggregated conversational recordings for some of the factors discussed in the literature, like task difficulty and emotional state, would be much too speculative to be useful. Overall, it appears that experimental methods are truly necessary to narrow in on this lawful relationship.

Yet, all is not lost. In this chapter, I show that this "lawful relationship" and our inability to adequately model it in a post hoc fashion can be leveraged to better understand the nature of sociolinguistic variation. The variability of pauses – and HESITANCY in speech timing more generally – will shed light on the processes of sociolinguistic variable production. To explore this, we start by considering intra-speaker variation, or STYLE, in the sociolinguistic endeavor.

Integral to Labov's earliest discussions of style – and analyses of linguistic variation – was a realization that systematic means were necessary to assess with any rigor the actual degree of attention a speaker might be paying to her or his speech. Consequently, the sociolinguistic interview was designed to elicit a range of ATTENTION TO SPEECH styles through varying tasks that might draw a speaker's attention more, and more importantly less, to the act of speaking. Labov (1966[2006], 1972) demonstrated the value of eliciting speech styles based on tasks that range from (most formal/attentive) readings of word lists and prepared reading passages to conversational interview speech (considered careful speech) and highly emotive or naturalistic speech (considered less careful speech; least formal/attentive). Labov showed, for example, that prompting interviewees to discuss "danger of death" stories (e.g. "Have you ever been in a situation where you thought you were in serious danger of being killed – where you thought to yourself, 'This is it?'") often elicited more casual talk than other, formal interview prompts (Labov 1972: 92–4).[1]

At the same time, to address the question of just how we might be sure that these differing tasks succeed in eliciting different styles, it was necessary to look to other aspects of the sociolinguistic interview data (i.e. other aspects of the speech recordings). It is a circular argument to use linguistic variable data (such as decreasing rates of *r*-fulness in New York City as tasks become increasingly informal) to argue that speakers undergo changes in attentiveness and at the same time that changes in attentiveness impact speakers' variable productions. To "solve" this problem, Labov introduced the idea of CHANNEL CUES, or what have occasionally been termed PARALINGUISTIC CUES:

> It is of course not enough to set a particular context in order to observe casual speech. We also look for some evidence in the type of linguistic production that the speaker is using a speech style that contrasts with Style B [i.e. careful speech]. To use phonological variables would involve a circular argument, because the values of these variables in Styles A and B are exactly what we are trying to determine by the isolation of styles. The best cues are channel cues: modulations of the voice production which affect speech as a whole. Our use of this evidence must follow the general procedure of linguistic analysis: the absolute values of tempo, pitch, volume, and breathing may be irrelevant, but contrasting values of these characteristics are cues to a differentiation of Style A and Style B. A *change* in tempo, a *change* in pitch range, a *change* in volume or rate of breathing, form socially significant signs of shift towards a more spontaneous or more casual style of speech. (Labov 1972: 94–5, emphasis in original)

In sum, Labov proposed the use of channel cues as a systematic means to assess whether (and when) the interview-based strategies he developed were successful in manipulating speakers' attention to their speech.

It is notable that at the same time that Labov was developing the variationist sociolinguistic program, his attention to speech model, and the idea that channel cues might provide valuable windows into speakers' (changing) attention to their speech, Frieda Goldman-Eisler and other psycholinguists were developing their productive line of psycholinguistic research that approached many of these same channel cue features – e.g. pauses and speech rates – from the perspective of what they tell us about processes of speech production (see Chapter 2). The parallels between these lines of research are striking even though their connections appear to have been mostly ignored for the past 40 years.

In this chapter, we reconsider channel cues from the perspective that speech timing phenomena, and pauses in particular, provide a window into the processes underlying speech production. First, in §8.2 and §8.3, we continue to discuss STYLE, as intra-speaker variability is most often described in sociolinguistics, and consider the relationships between sociolinguistic conceptions of style and related areas of psycholinguistic research. Then, in §8.4 and §8.5, we draw on work from Frieda Goldman-Eisler and her colleagues (Henderson et al. 1966) to develop a graphical and analytic method called the HENDERSON GRAPH. This technique provides quantitative metrics for assessing channel cues to intra-speaker variation, which, in §8.6, we apply to a case study analysis of language variation and variable (ing) in the speech of the African American adolescent girls from Washington, DC, whom we have seen previously in §7.3.2 (see also Mallinson and Kendall 2009, Kendall 2010b). Finally, §8.7 sums up and closes this consideration.

8.2 The sociolinguistics of style

As mentioned above, the stylistic dimension of language variation has been a central area of sociolinguistics since its formal foundations. And it has, of course, extended to multiple areas beyond Labov's attention to speech model. Most reviews of sociolinguistic approaches to style (cf. Schilling-Estes 2002) organize these approaches as falling into three main categories. I briefly present these here – following the general practice – in chronological order. (Many good reviews of style exist in the sociolinguistic literature, and I do not present a full review here; see Schilling-Estes 2002 for a thorough review of these approaches and Eckert and Rickford 2001 for an excellent and diverse set of discussions.)

As was outlined above, Labov introduced the quantitative study of sociolinguistic style by equating "style" with a speaker's attention to his or her speech. This operationalization, termed ATTENTION TO SPEECH, was intended as a systematic means to categorize the range of speech types elicited in sociolinguistic interviews, but it quickly became the de facto theoretical approach to style used by quantitative sociolinguists.[2] Importantly – especially for the present discussion – Labov did not, in fact, intend for his attention to speech approach to be used as an all-encompassing theory of speaker style. Instead, he viewed this idea as an important technique for the analysis of sociolinguistic interview data. As he later explained:

> The fact that these four or five styles can be ordered by increasing attention paid to speech has been mistaken for a claim that this is the way that styles and registers are to be ordered and understood in everyday life. The style shifting devices used ... were introduced as heuristic devices to obtain a range of behaviors within the individual interview, not as a general theory of style shifting. (Labov 1966[2006]: 59)[3]

It was not until the 1980s that alternative approaches to style were seriously considered in variationist work. Bell (1984; see also Rickford and McNair-Knox 1994, Bell 2001) brought the earlier social psychological models of style as accommodation based on inter-speaker relationships (Giles 1973, Giles and Powesland 1975, Giles, Coupland, and Coupland 1991) to quantitative, variationist sociolinguistic practice for the first time with his introduction of AUDIENCE DESIGN as a model of speech style. Audience design moved the primary focus of "style" from a speaker's attention to her or his speech to aspects of the speaker's audience. While this led the way to the testing of new hypotheses about language variation and language and identity and so forth – such as Bell's "Style Axiom" (1984: 151) – it has also been criticized for maintaining the same sort of unidimensionality that characterizes the attention to speech model. Although amended somewhat in Bell (2001), the audience design model is similar to the attention to speech model in that it frames all intra-speaker variation as derived from a single factor (or set of factors) when it is clear (through research like Eckert 2000 and Mendoza-Denton 2008) that speakers both respond to a more multiplex range of factors and initiate stylistic shifts for a range of reasons (see also Schilling-Estes 2002, Coupland 2007).

The most current set of approaches to understanding stylistic variation falls into what Schilling-Estes (2002) terms SPEAKER DESIGN APPROACHES. These approaches view style "not as a reactive phenomenon but as a resource in the active creation, presentation, and recreation of speaker identity" (2002: 388). Speaker design approaches have led to a number of advances in our understanding of language and identity, speaker agency, and so forth (see e.g. Podesva 2007). However, as Schilling-Estes (2002: 392–4) explains, they also raise some problems for variationist linguistics that were minimized by the earlier approaches to style. Speaker design approaches have tended to be largely qualitative (as opposed to quantitative), interpretative (as opposed to predictive), and microsocial or individualistic (as opposed to macrosocial). There is, of course, nothing wrong with this – and, in fact, there are some substantial benefits to this – but it has been hard to integrate into the firmly quantitative, predictive, and somewhat macrosocial foci of the main body of variationist work, with its overarching goal of understanding the principles behind language variation and change (cf. Weinreich et al. 1968, Labov 1994, 2001, 2010). If possible, extending speaker design approaches to speak more to broader variationist inquiries would benefit many areas of language study.

In very recent work, Eckert (under review) articulates an approach to studying variation that places a central focus on STYLISTIC PRACTICE. She writes, "stylistic practice takes place in moment-to-moment adjustments, as stylistic agents encounter styles that they interpret as standing in some important relation to their own" (ms, p. 20). This work – and that of her colleagues in the THIRD WAVE (e.g. Podesva 2007, Zhang 2008) – points towards a way forward for enriching the treatment of style in sociolinguistics. The unidimensionality of the earlier approaches to style seems to be a thing of the past.

So, is attention to speech, as an approach to style, dead? Even considering style from the most multiplex of recent perspectives we must acknowledge, quoting Bell, that "there is undoubtedly some relationship between attention and style" (1984: 150). Bell goes on to say that, in fact, "attention is at most a mechanism of response between a situation and a style." What I argue in this chapter is the idea that attention to speech is not "style" in the more complete sense that we now wish to give style, but that it is nonetheless a very important phenomenon that, when considered in its true light, can inform sociolinguistics even more than it already has. Put simply, attention to speech is a cognitive, not a stylistic, phenomenon. And through it – as we have seen since Labov's New York City study (1966[2006]) – we can learn about the processes underlying language variation as it is produced in natural speech.

A key to moving forward with the development of sociolinguistic theories and methods for the understanding of stylistic variation is to explore ways that we might integrate the correlational power of the attention to speech model with more sophisticated and nuanced understandings of language and identity (e.g. Schilling-Estes 1998, Coupland 2007, Mendoza-Denton 2008, Eckert 2000, 2005, under review). This becomes possible, I believe, when we consider individual variation from a psycholinguistic perspective.

8.3 The psycholinguistics of style

At about the same time that William Labov was laying the foundations for the modern field of (variationist) sociolinguistics, Frieda Goldman-Eisler was developing her psycholinguistic research on sequential temporal patterns of talk (e.g. pause and speech rate) and investigating their relationships to speakers' language production processes.[4] As we have already seen, her work showed that much of spontaneous speech is "a highly fragmented and discontinuous activity" (1968: 31), that pauses are more likely and longer before words with less predictability and with more difficult speaking tasks, and that – in the terminology and conception of the time – pauses can be used "to sort out which parts of verbal sequences are verbal habits and which are being created at the time of speaking" (1968: 43). Further, Goldman-Eisler explained:

> Pausing during the act of generating spontaneous speech is a highly variable phenomenon which is symptomatic of individual differences, sensitive to the pressure of social interaction and to the requirements of verbal tasks and diminishing with learning, i.e. with the reduction in the spontaneity of the process. (1968: 15)

Many of Goldman-Eisler's conclusions have been paralleled by other psycholinguists who have pursued questions of speech timing. For example, Maclay and Osgood's early (1959) work found that hesitation pauses are more often realized before a semantically heavy unit than at clause boundaries (see also recent work like Arnold, Fagnano, and Tanenhaus 2003, Arnold, Hudson Kam, and Tanenhaus 2007). In general, Goldman-Eisler's major findings appear to have been confirmed numerous times and in numerous ways (e.g. Kircher, Brammer, Levelt, Bartels, and McGuire 2004; see, more generally, Levelt 1989). And, as we have discussed throughout this work, other researchers have examined the ways that one can map the "lawful relationship between temporal phenomena in human speech and concurrent cognitive processes" (Kowal and O'Connell 1980: 61).

Attention to speech and its paralinguistic cues have also had parallels in other areas of linguistic work. For example, in Chapter 2 I mentioned Wallace Chafe's work on the *Pear Stories* (Chafe 1980a), in which Chafe used pauses to help better understand information flow in discourse. In later work, Chafe (1994) introduced an approach to considering discourse that has many similarities to Labov's attention to speech. Instead of "attention" however, this approach centers on humans' "consciousness" and the ways that CONSCIOUSNESS unfolds in and drives discourse.

> Consciousness is the activation of only a small part of the experiencer's model of the surrounding world, not the model in its totality. This limited capacity of consciousness is reflected linguistically in the brief spurts of language that will be discussed as *intonation units* [later in his book]. Each such unit verbalizes a small amount of information which, it is plausible to suppose, is that part of the speaker's model of reality on which his or her consciousness is focused at that moment. In a socially interactive situation it is the portion on which the speaker intends that the listener's consciousness be focused as a result of hearing the intonation unit. (Chafe 1994: 29)

In sum, talk arises in spurts and these spurts parallel a speaker's conscious attention. Chafe's interest in his 1994 book centers on questions of information structure in discourse, and the bulk of the work examines talk and writing from this perspective, arguing for a number of principles based on notions of FLOW and DISPLACEMENT of consciousness. Considered even loosely, however, his treatment of consciousness could readily be extended to questions regarding social meaning and sociolinguistic style. In fact, consciousness, as implemented by Chafe and illustrated in the above quotation, appears to nicely integrate the range of sociolinguistic treatments of style (attention to speech, audience design, and speaker design).

Chafe's interest in information structure has further parallels in areas of psycholinguistics, and some related questions in particular have shed light on linguistic processes that we might consider related to sociolinguistic style and identity production. For example, Arnold (2008) reviews psycholinguistic and psychological research on the degree to which the production of reference (e.g. *her* versus *the girl*, and *the brown car* versus *the car*) is mediated by speaker-oriented and audience-oriented processes. This larger (and quite large) literature on speaker-oriented versus listener-oriented processes in language production seems quite relevant for the discussion of style and language and identity. Put together, these sometimes interwoven and sometimes independent lines of psychological and psycholinguistic research bear remarkable similarities to core sociolinguistic understandings of style. The extensive research on pause shows close parallels to Labov's concept of channel cues, and work such as that just mentioned by Arnold (2008) has parallels in sociolinguistic instantiations of audience design (i.e. Bell 1984, 2001) and broader communication accommodation theory (CAT; Giles et al. 1991). The consideration of bridges between these various lines of inquiry seems strikingly overdue. I attempt now to draft some bridges by focusing on the place of pauses in particular and speech timing phenomena more generally in the sociolinguistic implementation of attention to speech.

8.4 Channel cues to attention to speech

Despite the fact that Labov (1972) described a systematic approach to determining speakers' attention to their speech using paralinguistic cues, such as pauses, breathing, and speech rate, the complex relationship between sociolinguistic variables and speech features such as pause and speech rate has rarely been examined in a systematic, quantitative way. Labov's own focus on these features appeared to taper off following his original explorations of paralinguistic cues and he later wrote: "it appears that channel cues did not provide a high enough level of interpersonal reliability for most researchers" (1966[2006]: 74). And it does seem to be the case that other researchers failed to find Labov's channel cues actionable or useful. For example, in his groundbreaking study of African American English in Detroit, Wolfram (1969: 58–9) noted:

> An exploratory attempt to distinguish careful from casual speech based on Labov's criteria was rejected for several reasons. In the first place, any of the paralinguistic channel cues cited as indications of casual speech can also be indications that the informant feels an increased awareness of the artificiality or formality of the interview situation. Can nervous laughter reliably be distinguished from casual or relaxed laughter? Also, the subjective interpretation of the paralinguistic cues tends to bias the interpretation of casual speech even though the channel cues are theoretically supposed to be independent of the measurement of linguistic variables. To what extent must there be a change of pitch or rhythm and how close to the actual feature being tabulated must it occur?

Very little variationist work following these early studies has followed up on the possibility that speech features like speech rate and pause are useful indicators into sociolinguistic variation. The main exception to this falls in studies which have looked at pause as a predictive phonological environment in the patterning of certain sociolinguistic variables, such as SYLLABLE-CODA CONSONANT CLUSTER REDUCTION, also called FINAL STOP DELETION, T/D DELETION, or, most recently, CORONAL STOP DELETION (CSD; Guy and Cutler 2011, Hazen 2011). Numerous studies (e.g. Guy 1980) have found following pause to be an important environment in the patterning of CSD. In other words, analysts such as Guy have coded for when the following environment is "quiet" (Guy 1980's code "Q"). In some cases, such as Fasold's (1972) study of African American English in Washington, DC and Wolfram, Childs, and Torbert's (2000) study of the variable in African American English in Hyde County, pause appears to pattern along with consonants for this variable, but in others (e.g. the Cherokee Sound Anglos discussed in Wolfram et al. 2000) pause patterns more similarly to vowels. In Guy's own (1980) analysis, pause emerges as a factor that differentiates the white dialects in New York City versus Philadelphia.

In this usage, however, pause is treated as a structural category – it is the absence of a segment in the following position. CSD studies have not considered the importance of pause length or other, more suprasegmental factors. Several researchers have, however, indicated that rate of speech is likely an influence on CSD rates. For instance, Guy noted that "probability of deletion apparently increases in proportion to the rate of speech" (1980: 9). At the time of Guy's 1980 paper, there was not yet a reliable way to measure rate or speech, so he did not include it in his analysis.[5] It is this sort of more "paralinguistic" assessment of both pause and speech rate that we turn to and pursue further now.

As discussed in the previous section and elsewhere in this book, pause has been found to correlate with cognitive processing (Goldman-Eisler

1968) and, as demonstrated by Chafe (1994), discourse can usefully be segmented and analyzed in "spurt-like" increments. Understanding attention to speech and pausing as not only related to one another but also related to more general cognitive processes (such as processes of language production and identity production) may lead to a richer conception of linguistic variation. Considering channel cues as psycholinguistic variables further links the concept to a large literature where we may be able to draw on methods, in addition to understandings, that could be productively incorporated into sociolinguistic analysis.

Thus, I argue that a new focus on attention to speech and, in particular, on Labov's channel cues in cognitive terms will enable new insights into language variation. Importantly, this does not necessitate that we forego the methodological and theoretical advances that speaker design approaches to style (Schilling-Estes 2002, Coupland 2007, Eckert under review) have developed or, on the other hand, that we take up the sort of strictly correlational view of style that might come out of a straightforward reading of Labov's (1972) argument for attention to speech. In fact, it is instead to take up a point that Schilling-Estes makes in her review of approaches to style – that "examining as many types of features as possible is crucial in any research enterprise concerned with speaker meaning" (2002: 390). Considering STYLE here as something broad, as something bound up with online processes of identity management, as something both responsive and initiative (cf. Bell 2001, Schilling-Estes 2002) but as situated in the dynamics of language production and interpersonal interactions, we can rigorously consider the relationship between psycholinguistic processes and speaker meaning. However we define "style," there are cognitive processes at work that underlie speakers' linguistic and identity-related productions, and by better understanding those processes and integrating their examination into our sociolinguistic analyses we can open up paths to richer and deeper understandings of stylistic variation and language and identity.

8.5 The Henderson graph: a method for quantifying attention to speech

As the previous chapters of this book have demonstrated to some length, the earlier problems (such as those discussed by Guy 1980) in accurately measuring rate of speech and other temporal and paralinguistic phenomena no longer present such barriers. Through instrumental techniques, the mechanics of measuring rate of speech and pause are relatively

straightforward. And, as the previous chapters have demonstrated, with accurately time-aligned transcripts available, corpus sociophonetic methods can be used to rapidly generate a huge amount of speech timing data. Nonetheless, asking how speech rate and pause may interact with, affect, or predict variable production is not at all a clear issue. For example, in attempting to quantitatively map the realization of pauses against the realization of a particular variant form, should one measure the *duration of* the most recent pause, or the *duration since* the most recent pause, or both? Should one attend to the *number* of pauses more than (or as much as) their *durations*? We are reminded of Wolfram's early critique: "To what extent must there be a change of pitch or rhythm and how close to the actual feature being tabulated must it occur?" (Wolfram 1969: 59).

In searching for these sorts of possible correlations how do we know when or if we have discovered the most meaningful relationship? And, how do we know when or if we have found merely a symptom of the linguistically or cognitively meaningful relationship instead of the relationship itself? I do not claim fully to answer these difficult questions here but I propose an approach to investigate them. To do this, I adopt a broader-based metric than simply pause or speech rate. I here (re-)introduce the HENDERSON GRAPH, a technique piloted in the psycholinguistic research on speech timing of the 1960s (Henderson et al. 1966, Goldman-Eisler 1968; see also Levelt 1989, Kendall 2009, Thomas 2011a). The Henderson graph, demonstrated in Figure 8.1, is a representation of a speech event in which talk time is plotted on the *x*-axis while pause time extends along the *y*-axis. Stretches of talk are represented in a stair-like line plot extending over units of time (displayed as seconds in the example of Figure 8.1). Changes in the characterization of the talk are viewable as changes in the slopes of sections of the graph and these slopes and changes in slope can readily be measured, as illustrated in the figure. Talk that is more hesitant – that is, that has a greater proportion of pause time to talk time – extends vertically more quickly than horizontally. Stretches of talk can be compared via their slopes to one another, while at the same time overall or average slopes can be computed for speakers and/or interviews and can be used to make broader comparisons across speakers, interviews, and so forth (Henderson et al. 1966; cf. Thomas 2011a).

In their original work developing the Henderson graph, Henderson et al. demonstrated that talk is generally characterized by alternations in slope lines between steeper and shallower slopes and that steeper slopes were not only characteristic of a greater proportion of silent pauses, but also had significantly more filled pauses, false starts, and nongrammatical

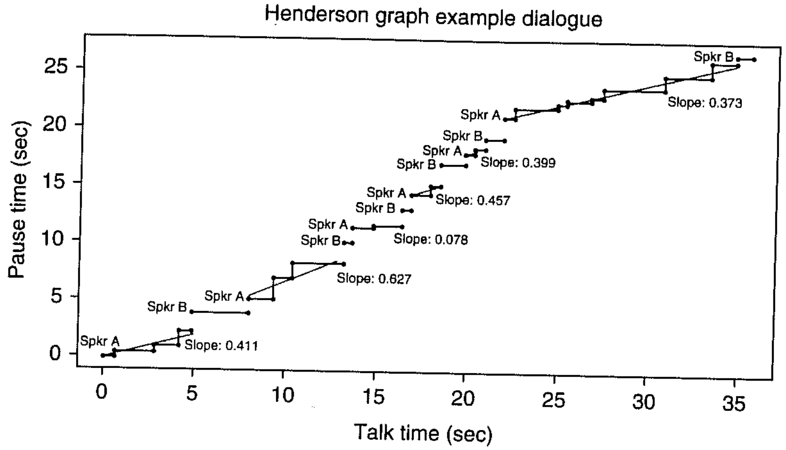

Figure 8.1 Example of a Henderson graph for an interview dyad

"gaps" (i.e. hesitation pauses, not simply pauses between clauses or syntactic units) than shallower slopes (1966). Henderson et al. showed that the graphing technique captured more information about language production processes than simply the proportion of talk time to silent pause time which it mechanistically reports (in purely mathematical terms the slope provides a smoothed ratio of pause-to-talk time). Further, they demonstrated that there is "a regular non-random distribution of hesitation phenomena throughout spontaneous speech" (1966: 214) and that this nonrandom distribution could be captured and used to segment talk into planning units and execution units "based on what could be called cognitive rhythm or cognitive stride" (1966: 216).[6]

Henderson et al. focused on sections of monologue in spontaneous speech. Their data came from interviews, but for their analysis they extracted longer stretches of uninterrupted talk by a single speaker. In conversational speech, and the sort of talk contained in many sociolinguistic interview recordings, many talk-turns are quite short and it is very common for speakers' to have turns that are comprised of only a single phonetic utterance. These single-utterance turns are realized in Henderson graphs as horizontal lines, as illustrated by Spkr B's contributions in Figure 8.1. These segments have a slope of 0. In the treatment here (i.e. in the case study below), these segments are excluded from analysis and only stretches of talk containing at least one turn-internal pause are included in the analysis (although further evaluation would

usefully shed light on the best way to treat these segments). Also, following Henderson et al. (1966), filled pauses like "uh" and "um" are included in the phonetic utterances of talk. Other techniques and modifications could be developed to better account for these kinds of features, but we stick to as simple an implementation as possible for sake of exploring the potential benefit of the graphing technique.

As we will explore, the slope measurements provide a quantitative metric that can be correlated with speakers' variable productions. Additionally, the HENDERSON SEGMENTS – the stretches of talk by a speaker that are fitted to a single slope line – provide an empirical segmentation of the discourse. In Figure 8.1, each Henderson segment is equivalent to a speaker turn. However, longer stretches of talk by a single speaker are divided into multiple segments. Again, as Henderson et al. (1966) investigated in depth, these longer stretches are generally characterized by an alternating pattern of shallower and steeper slopes. Changes in the Henderson graph slope lines indicate points at which to segment the discourse into smaller "chunks." The Henderson segments allow us to test other speech timing features in addition to slope (like pause duration, speech rate, and even length of segment) against variable realizations. This approach, I argue, provides an empirical and organic means with which to segment talk. Otherwise, if we want to assess, say, the potential influence of articulation rate on a variable realization we are stuck with Wolfram's (1969: 59) problem of having to make an arbitrary or subjective decision about what unit of speech is relevant for a given variable token. The consideration of how best to "chunk up" conversational speech is – and will likely always be – up for argument. The approach to segmenting the talk into units larger than the phonetic utterance using the changes in slope within the Henderson graph seems to me to be a good starting point. Future research will, I hope, be productive for refining these procedures.

8.5.1 A new methodology for Henderson graphing

While Henderson et al. (1966) did not go into extensive detail into the specifics of the generation of their sequential temporal graphs, one must imagine that it was a painstaking and slow process when compared to modern software tools for analyzing and visualizing acoustic data. Current software (such as Praat; Boersma and Weenink 2010) makes the delimitation of pause and speech fairly straightforward, and we can even design software to automatically generate the Henderson graphs for data of interest once we have the temporally delimited data. In order to generate Henderson graphs for arbitrary stretches of audio-recorded talk, I have developed a software-based feature in SLAAP (see Chapter 3 for an introduction to SLAAP and its transcription archive). This is realized as a web page which takes the fine-grained time-aligned transcripts in the SLAAP archive and presents them dynamically in a variety of formats (Kendall 2006-2007, 2007a), including a Henderson graph version. An example of a SLAAP-generated Henderson graph is shown in Figure 8.2. This sample presents a part of a single speaker's monologue, which in total consisted of about one minute of talk by the speaker with no interruptions or contributions from other speakers. In truth, such long stretches of uninterrupted talk are rare in the conversational, interview recordings used throughout this book, and most speaker turns are more like those depicted in the example of Figure 8.1.

The variable extraction and coding features in SLAAP automatically record accurate timestamp information for each variable tabulation when an analyst codes its contextual and structural factors (cf. Tagliamonte 2006 for a broad, detailed overview of variable analysis procedures and Wolfram 1993 for a shorter discussion of sociolinguistic variables and tabulation). The variable tabulation data and the transcripts are stored in SLAAP's database and, since both sets of data maintain highly accurate timestamp information, these data can readily be merged and extracted from SLAAP through a combination of SLAAP's web-based features and additional extraction scripts (for the work here, all SLAAP-external scripts were built in the R language; R Project Development Team 2011). Note that Figure 8.2 shows variable tags in situ, indicating where some coded variables occur in the graphed discourse.

The SLAAP Henderson graphing software also aids in the generation of slope lines by applying a best-fit linear model to each speaker turn containing at least one intra-turn pause. Many short talk-turns are appropriately segmented into a single Henderson segment, i.e. are best characterized by a single slope line. For longer stretches of talk, as in the SLAAP-generated Figure 8.2, the analyst can select points in the turn where the slope lines should break and then the software recomputes best-fit lines for each segmented section. Following Henderson et al., this work is necessarily interpretive and conducted by analyst "inspection" (1966: 208). Oftentimes a combination of the discourse content (e.g. audio content and transcript text) and character of the graph is useful for determining whether a given turn should be partitioned into multiple Henderson segments. This is, as it sounds, somewhat subjective work, and there is clearly room for improvement. In general, for the work presented here I have taken a conservative approach to the

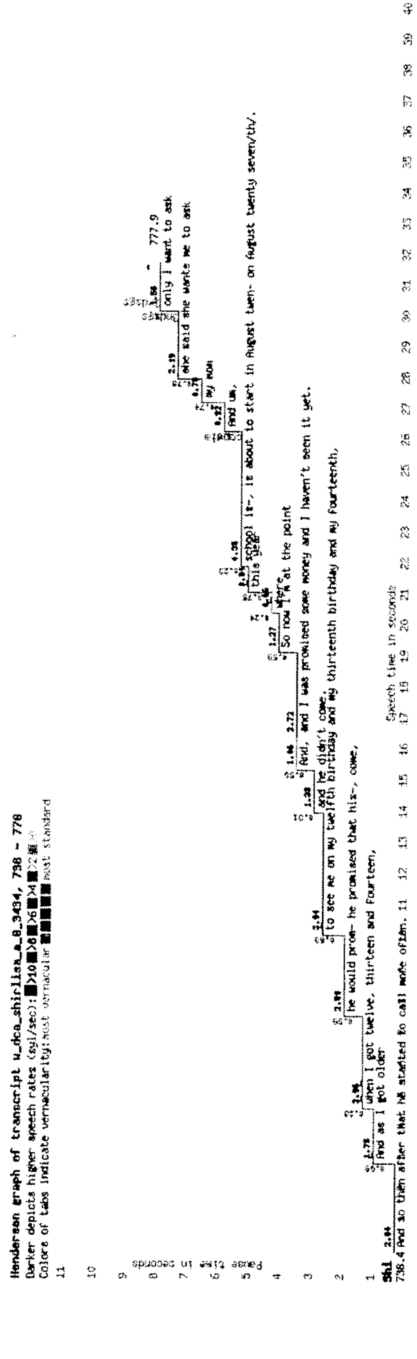

Figure 8.2 SLAAP screenshot of a Henderson graph

segmentation, only segmenting long turns when there is clear evidence for a change in the discourse. Thus, most Henderson segments correspond to speaker turns. As I hope to show, the current implementation, as described here and used for the upcoming case study, nonetheless demonstrates promising results.

8.5.2 Henderson graph-based metrics

Through the Henderson graph we can develop a wide range of quantitative variables for potential (socio)linguistic analysis. Table 8.1 provides a short enumeration of some possible measures that can readily be extracted from the Henderson graphs.

Other measures not listed here could also be generated from the Henderson graph method and Table 8.1 only illustrates some likely candidates of interest. The SLOPE, ΔSLOPE, SLOPECOMP, and R^2 variables are directly related to the slope of the best-fit line for a given Henderson segment. The other six variables listed, DUR, PAUSEN, PAUSEDUR, ARTRATE, ARTRATECOMP, and SPKRATE, are determined based on the stretch of time within the Henderson segment; they are not measures derived from the Henderson graph itself. These variables (e.g. ARTRATE and ARTRATECOMP) can be used to determine if features like rate of speech correlate with variable productions. While not derived directly from the Henderson graph, they are still dependent on the

Table 8.1 Some Henderson graph-based variables

SLOPE = Best-fit slope over the current Henderson segment[7]
ΔSLOPE = Change in slope from previous Henderson segment
SLOPECOMP = Tertiary variable based on the comparison of a given slope and that speaker's mean slope (low, norm, high)[8]
DUR = Overall duration of Henderson segment
PAUSEN = Number of pauses within segment
PAUSEDUR = Median pause duration within segment
ARTRATE = Median articulation rate within segment (σ/second, not including pauses)
ARTRATECOMP = Tertiary variable based on the comparison of a given segment's median articulation rate and that speaker's overall articulation rate (low, norm, high)
SPKRATE = Overall speaking rate for segment (total # σ/total duration of segment, includes pauses)
R^2 = Correlation coefficient for current slope line

graphs as they are measured in units based on the segmentation of the talk by its changes in slope.

My hope is that these Henderson graph-based variables will let us articulate and test new questions about language variation. We now examine some real data, in order to evaluate some of these variables and to assess some possible utilities of the Henderson graph technique for variationist linguistic analysis.

8.6 Case study: the interviews with adolescent African American girls in Washington, DC

In order to test of the utility of the Henderson graph we return to the ten interviews with the young African American females from Washington, DC (discussed in §7.3.2; Mallinson and Kendall 2009). This time we visit these interviews to ask: can we better understand sociolinguistic data and better model sociolinguistic variation by incorporating paralinguistic detail into variationist analyses? And, more specifically, do the Henderson graph-based metrics shed light on the realization of sociolinguistic variables?

8.6.1 Henderson graph slopes and sequential temporal variation

I begin by briefly considering the Henderson graph-based data from these interviews as data in their own right. What do the graphs and slope measurements tell us about the nature of the ten interviews? And, what sorts of individual variation do we find? Table 8.2 provides a summary of the slope data by speaker. While the slopes of individual Henderson

Table 8.2 Slope summary for DC speakers

Speaker	Measurable segments	Min.	1st qu.	Median	Mean	3rd qu.	Max.
Ala-14	277	0.008	0.128	0.203	0.299	0.330	4.461
Asi-12	207	0.012	0.195	0.346	0.538	0.574	6.763
Cal-13	148	0.005	0.129	0.258	0.467	0.608	5.588
Eli-14	165	0.024	0.178	0.317	0.549	0.516	7.461
Gra-13	221	0.027	0.204	0.271	0.379	0.379	3.582
Kei-15	296	0.010	0.263	0.426	0.517	0.677	2.412
Lat-17	257	0.028	0.181	0.314	0.500	0.615	3.929
Sha-14	340	0.022	0.195	0.323	0.430	0.549	2.582
Sha-12	160	0.017	0.159	0.376	0.594	0.667	5.296
Shi-14	350	0.020	0.122	0.201	0.244	0.283	2.585
Overall	2,421	0.005	0.163	0.284	0.432	0.512	7.461

segments should provide a systematic window into some paralinguistic aspects of the talk (ultimately, our primary interest here), we can also use speakers' or discourses' mean slopes as ways to characterize the overall sequential temporal structure of those speakers or discourse events (Henderson et al. 1966, Thomas 2011a). We start with this latter view.

The table provides a full summary of each speaker's slope data, including information about each speaker's range of slope values and mean and median. The "measurable segments" column provides the Henderson segment *N* for each talker. Only turns longer than one single utterance are included here. At least one pause is necessary for a turn to have a nonzero slope and, as I indicated above, I have excluded all zero slopes from the present investigation. The average slope across all ten speakers is 0.432. In simple terms, this means that for measurable Henderson segments the speakers have an average proportion of talk time over pause time of just over 2 to 1; for every second of talk, there is on average 0.432 seconds of pause (turn-internally). An ANOVA finds that the slope data are significantly different by speaker ($F(9, 2411) = 20.478$; $p < 0.000001$). Considering slope as a measure of hesitancy, we can observe that the most hesitant speakers, Sha-12 and Eli-14, have mean slopes of 0.594 and 0.549 respectively, while the least hesitant speakers, Shi-14 and Ala-14, have mean slopes of 0.244 and 0.299 respectively. The mean slopes are depicted graphically in Figure 8.3. Again, keep in mind that higher slope values indicate higher proportions of pause-to-talk time.

In earlier work (Mallinson and Kendall 2009), Christine Mallinson and I examined these recordings and explored differences among the individual girls' interviews.[9] In particular, we noted that two of the girls,

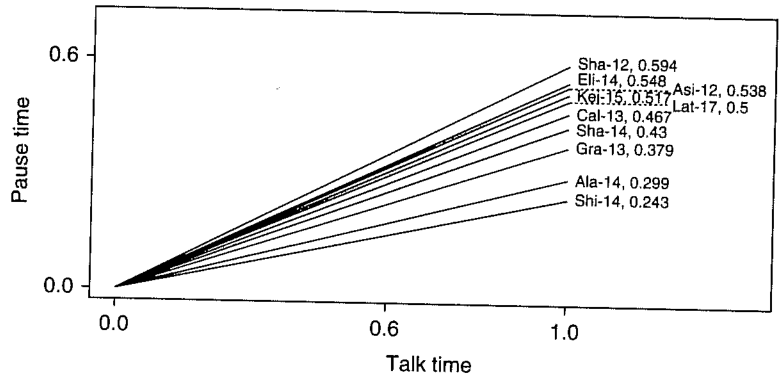

Figure 8.3 Mean slopes for DC speakers

Ala-14 and Shi-14, both realize the least vernacular speech and in many ways appear the most talkative and at ease in the interviews. When we consider the slope patterns of Table 8.2 and Figure 8.3 with respect to this preexisting understanding of the data, we find that the slopes fall in line with this sense of the speakers. Shi-14 and Ala-14 yield the shallowest average slopes. The quantitative Henderson graph slope metric successfully captures the larger discourse patterns in the talk, paralleling interpretations made through our other investigations. This finding across speakers appears to parallel Henderson et al.'s (1966) results within speakers, that the slope data correlate with broader signs of hesitancy.

8.6.2 Hesitancy in narrative versus nonnarrative talk

In a different paper, a talk presented at the Georgetown University Round Table on Language and Linguistics (Kendall and Mallinson 2008), Mallinson and I coded these interviews for narrative versus nonnarrative passages with the interest of determining the extent to which the spontaneous telling of narratives increased (or decreased) the likelihood of the realization of vernacular speech features in the interviews (Labov and Waletzky 1967, Labov 1972; see also Rickford and McNair-Knox 1994). For the entire dataset, stretches of talk were time-stamped and coded as "narratives" based on close readings of the transcripts and close listens to the audio. In general, narratives were: (a) in the past tense, (b) relatively uninterrupted by the interviewer, (c) begun with one of several discourse markers (e.g. "like," "like, okay," "lemme tell you"), (d) frequently closed with some sort of evaluative statement (e.g. "that's about it"; cf. Labov and Waletzky 1967), and (e) often about personal matters, like dating and family (Kendall and Mallinson 2008). Sometimes one narrative was embedded in another, and, overall, these narratives were often brought up as tangents to the interview talk – C, the interviewer, was not, for the most part, trying to elicit narratives. The following excerpt provides a short example of a narrative passage from Lat-17's interview.

(8.1) C: Does your mom drive? [gap 0.78 sec]

Lat-17: Sometime she drive, but [pause 0.20 sec] she, she about to get a new car, [pause 0.30 sec] cause my aunt was in an c- [pause 0.20 sec] accident. [pause 1.37 sec] That was like last year. [gap 0.06 sec]

C: Uh-huh.

Lat-17: /Tom/ and them were /?/ /and/ uh, [pause 0.45 sec] she was in an accident, and their car crashed, [pause 0.45 sec] but she was okay. [gap 0.75 sec]

C: And that was your mom's car? [gap 0.12 sec]
Lat-17: Yeah. [pause 1.52 sec] And now she got to get a new car.
(Lat-17 media file a: lines 546-572; time 384.48-408.21)

Labov's formulation of attention to speech indicated that engaged narratives (like "danger of death" stories) might be the sites for more *casual* speech than the *careful* speech thought to be obtained in much interview talk. The Henderson graph slope metric provides a means to test the hypothesis about narrative versus nonnarrative talk further. If, in general, the telling of narratives tends to be characterized by more casual, less self-conscious talk than other stretches of interview talk, we should expect that these passages would be less hesitant and that the slopes of the Henderson graphed talk would identify these sorts of conversational differences. We can ask: do narrative passages have a shallower slope than nonnarrative passages? To answer this I ran a mixed-effect regression testing the influence of narrative vs nonnarrative passage on slope (and including speaker as a random effect, since we saw above, for example in Figure 8.3, that individual speakers have different overall slope tendencies). This model yields a significant effect for the narrative status of passages ($p = 0.016$). Narrative passages are predicted to have slope values 0.0688 *shallower* than nonnarrative passages. The effect is by no means large, but it is important. Nonnarrative stretches of talk are significantly more hesitant by this metric than narrative stretches, even when we control for individual differences across speakers.

8.6.3 Attention to speech and variable (ing)

We now turn to consider the relationship between Henderson slope values and the realization of sociolinguistic variables. Variable (ing), the realization of the *-ing* ending as [ɪn], *-in'*, or [ɪŋ], *-ing*, in words like *talking, everything, pudding*, and so forth,[10] has been extensively studied in the sociolinguistic literature (e.g. Fischer 1958, Labov 1966[2006], Trudgill 1972, Tagliamonte 2004, Campbell-Kibler 2007, Hazen 2008, Kendall 2010b). (ing) has been consistently found to be a stable sociolinguistic variable, one which bears correlations with social features of speakers, like social class and education, and at the same time shows strong effects along the dimensions of speaker style, formality, and attention to speech (cf. Labov 2001, Kendall 2010b). For example, in a now classic, pre-Labovian study of language use in a New England village, John Fischer (1958) showed that the schoolchildren in his study

used the most amount of full *-ing* forms in the most formal settings and the least full *-ing* forms in the most informal settings.

(ing) provides a perfect variable with which to test the Henderson graph technique. If we expect that the rates of alveolar (ing) decrease (i.e. forms are realized with more fully velar, *-ing*, variants) when speakers are more attentive to their speech (Labov 1966[2006], 1972, 2001) and we expect that Henderson segments with steeper slopes indicate stretches of talk that are more hesitant (Henderson et al. 1966) then we can hypothesize that steeper-sloped Henderson segments should correlate with lower rates of (ing) fronting than do shallower-sloped segments. That is, Henderson slopes should be able to act as a quantitative and fine-grained proxy for something akin to attention to speech.

As a part of our larger project, Mallinson and I (Kendall and Mallinson 2008, Mallinson and Kendall 2009, Kendall 2010b) coded (ing) for all ten interviews. Unlike many variationist studies which tabulate variable rates based on quotas per speaker and use type-token limits to control for the fact that some words favor or disfavor rates of variables more than others, these (ing) data were extracted and coded for the entirety of the recorded interviews. No type-token limits were imposed. Mixed-effect modeling helps account for the different influences and rates of occurrence of different lexical items and ameliorates the need to limit the data to a certain number of tokens per word type per speaker.

The ten interviewees exhibit, in the aggregate, 83.4 percent (ing) fronting ($N = 1256/1506$). This is a high number, but high rates of *-in'* use are common in African American English (Labov 2001, Kendall 2010b). A simple mixed-effect regression model for these data – not yet including the Henderson graph-based factors – finds that grammatical category (GRAMCAT), the number of syllables in the *-ing* word (WORDSYLS) and following environment (FOLLENV) are significant factors in the realization of (ing) fronting.[11] Unlike the models presented earlier, the models here use mixed-effect logistic regression (also called generalized linear mixed-effect modeling) since the dependent variable is categorical rather than continuous. Modeling is similar to that described earlier for the mixed-effect linear models in Chapters 4 and 5. Instead of providing an estimated value for the dependent variable, however, logistic regression models determine a probability – expressed in log-odds – of a particular form being realized. Positive log-odds values favor the effect – i.e. favor the alveolar form – while negative log-odds disfavor the effect. Log-odds values further from zero indicate stronger effects. Also unlike the linear models, which were evaluated earlier by testing the correlation of the predictions with the actual data and generating R^2 values, we here use

Somers' D_{xy} and the C index of concordance statistics to assess model fit (Baayen 2008, Harrell 2009). Like R^2, these values fall between 0 and 1 and values close to 1 indicate good fits. Finally, in order to account for the fact that different words may have different tendencies for (ing) I treat the words as a random effect in the mixed-effect regression of the data in addition to including a speaker random effect. The best model is shown in Table 8.3.

Grammatical category is typically found to be an important factor in (ing) realization, with progressive verbal forms (e.g. "he is walking") realizing the highest rates of *–in'* (cf. Tagliamonte 2004, Hazen 2008, Kendall 2010b). For the model here, I have simplified the coding to just three categories: progressive verbs (N = 792), nouns and adjectives (N = 459), and gerundial forms (including gerundial participles, e.g. "I got here fast by running"; N = 255), with verbs as the baseline factor. Other scholars (cf. Tagliamonte 2004, Hazen 2008) have examined the influence of grammatical factors in more detail, but for our purposes this simple three-level category is adequate. The verb category is set as the baseline for the GramCat factor group. As we see from Table 8.3, the difference between verbs and nouns/adjectives is quite significant, with the latter category disfavoring fronting with a log-odds of –1.030 (p = 0.0064). The gerundial category also trends in this same direction, but the effect is not significant (p = 0.1622). The disfavoring status of nominal forms is in line with previous findings (cf. Tagliamonte 2004).

Following phonological environment is found in these data to be predictive of (ing) fronting, even though the data were only coded for alveolar forms (N = 351) and velar forms (N = 26), with all other environments coded as "other" (N = 1129). Alveolar forms favor fronting (an assimilation effect). Alveolar consonants are the baseline, so the model reports that both following velars (log-odds –0.716, p = 0.0090)

Table 8.3 Basic mixed-effects regression model for DC (ing) data

Factor	Log-odds est.	Std err.	Z value	p
(Intercept)	5.107	0.965	5.293	–
GramCat = noun and adj.	–1.030	0.378	–2.727	0.0064
GramCat = gerund	–0.356	0.255	–1.398	[0.1622]
WordSyls	–0.821	0.363	–2.259	0.0239
FollEnv = velar	–0.716	0.274	–2.614	0.0090
FollEnv = other	–1.603	0.666	–2.406	0.0161

Somers' D_{xy} = 0.842; C = 0.921.

and other following environments (log-odds –1.603, p = 0.0161) disfavor *–in'*. Although the data are sparsely coded for a full phonological analysis – this is not really our purpose at present – the finding that alveolar following consonants favor (ing) fronting more than velar following consonants is in line with Tagliamonte's (2004) analysis. Finally, the number of syllables (WordSyls) is also significant, showing that longer words are less likely to have *–in'* than shorter words. The model's estimate shows a log-odds *decrease* of 0.821 per syllable (p = 0.0239). In sum, even with their high rates of (ing) fronting, the DC speakers' patterns for (ing) are quite in line with typical findings for the variable.

Our main interest here is in the value of the Henderson graph-based predictors. Do any of these – in particular the slope-related predictors – aid in our ability to model the (ing) data? Following the basic principles laid out earlier in Chapters 4 and 5, numerous mixed-effect models were built testing the influence of the Henderson graph variables outlined in Table 8.1. The best model for the (ing) data was found when the model from Table 8.3 was expanded to include two of the Henderson graph variables. clSlope, the log-transformed and centered Henderson Slope, is highly significant. The articulation rate of the Henderson segment, seen through the tertiary predictor ArtRateComp, is also significant. This fuller, best model is shown in Table 8.4 and its effects are plotted in Figure 8.4.

For the effects viewed earlier, GramCat, FollEnv, and WordSyls, we obtain roughly similar results and, for sake of space, I will not comment on them further. We are primarily interested in the role of the Henderson graph-based predictors, clSlope and ArtRateComp. clSlope, the log-transformed and centered (within speaker) Slope value, shows that when

Table 8.4 Full mixed-effects regression model for DC (ing) data

Factor	Log-odds est.	Std err.	Z value	p
(Intercept)	4.920	0.972	5.063	–
GramCat = noun and adj.	–1.046	0.374	–2.798	0.0051
GramCat = gerund	–0.271	0.262	–1.033	[0.3015]
WordSyls	–0.800	0.361	–2.216	0.0267
FollEnv = velar	–0.732	0.275	–2.666	0.0077
FollEnv = other	–1.676	0.669	–2.505	0.0123
clSlope	–0.468	0.162	–2.888	0.0039
ArtRateComp = high	0.645	0.233	2.770	0.0056
ArtRateComp = low	0.082	0.245	0.335	[0.7377]

Somers' D_{xy} = 0.844; C = 0.922.

speakers' slopes increase, the likelihood of *–in'* decreases (log-odds –0.468 per unit of CLSLOPE, $p = 0.0039$). This falls nicely in line with the prediction that increased hesitancy – increased attention – on the part of a speaker leads to greater full velar *–ing* use. This is very exciting.

It is worth commenting more on the transformation of slope to this centered factor. By centering this factor on a per-speaker basis, this factor becomes a measure of how similar each Henderson segment's slope is to its speaker's mean slope. As such, the centered version of the factor provides a speaker-specific, relative measure of hesitancy, which is independent of the fact that each speaker has her own general slope tendency (as seen in Figure 8.3). In other words, centering here NORMALIZES the slope values across speakers. Many versions of the slope predictor were tested in the statistical analysis (all of those listed in Table 8.1 and some) and this logged and centered version was found to be the best predictor for (ing). This indicates, as we should expect I think, that it is not the absolute value of slope (the actual pause-to-talk ratio) that matters but rather a speaker's divergence from her or his typical slope. In a sense, this CLSLOPE version of the variable captures something between the actual, continuous SLOPE variable of Table 8.1 and the SLOPECOMP tertiary categorical variable proposed there.

We also find an effect for articulation rate, indicating that faster talk in purely articulatory terms (as articulation rates do not include pauses) also corresponds to higher rates of *–in'* use. While the continuous CLSLOPE predictor outperformed the SLOPECOMP factor in preliminary models of the (ing) data, the ARTRATECOMP tertiary variable surfaced as the articulation rate factor that best contributed to the success of the modeling. Preliminary models did find that actual articulation rate values arose as significant but those models performed less well than the model presented here. It is perhaps surprising that this effect is best captured by the categorical version of the predictor, but the actual results for ARTRATECOMP give some indication for why this predictor performs best in categorical terms. Articulation rates classified as "high" (more than a half standard deviation above the speaker's mean) are significantly more likely to occur with *–in'* (log-odds 0.645; $p = 0.0056$) than those in the speaker's normal range. Stretches of talk with slower rates, however, are not significantly different than those in the normal range. In other words, it is only the fastest utterances that show this articulation rate effect.

Altogether the findings here are highly interpretable and are, I think, very promising. Speech timing is an important factor in the realization of (ing). The model without the Henderson graph predictors (in Table 8.3)

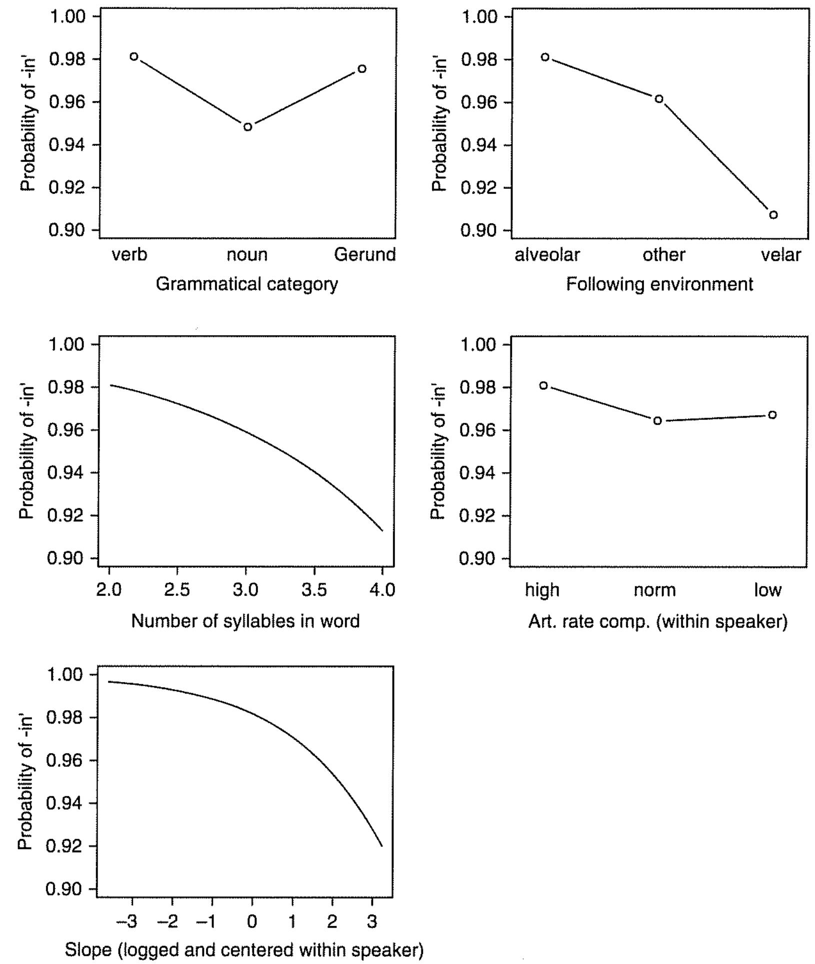

Figure 8.4 Effect from mixed-effect model for DC (ing)

obtained a good model fit, with Somers' $D_{xy} = 0.842$ and $C = 0.921$. The full model, with CLSLOPE and ARTRATECOMP, did better, but, we note, only by a small amount, with Somers' $D_{xy} = 0.844$ and $C = 0.922$. However, testing the results of a model built with only the CLSLOPE factor (although still with random intercepts for speaker and word), we remarkably obtain an even better fit, with Somers' $D_{xy} = 0.849$ and $C = 0.925$. Slope is an extremely strong predictor of (ing) realization.

8.6.4 Channel cues in the DC interviews

This case study has explored some of the possible uses of Henderson graphs. It has demonstrated that the Henderson graph-based metrics can be used as measures of paralinguistic cues to speaker style. For the data examined here, these measures have successfully captured paralinguistic differences between individual speakers and interviews and between narrative and nonnarrative stretches of talk within the individual interviews, helping to separate careful from casual speech (Labov 1972). Perhaps most importantly, these measures can be used as independent variables in statistical models of the sociolinguistic data. As demonstrated in the brief analysis of variable (ing), the inclusion of these paralinguistic predictors can improve our statistical model of the linguistic data and, I believe for one of the first times, quantitatively shows that there is indeed a direct relationship between speaker hesitancy and the realization of variable (ing). By doing this, we shed new light into the cognitive basis of language variation.

8.7 Conclusion

This brief case study and the larger explication of the Henderson graph technique have attempted to link variability in speech timing – and in particular pause realization – to some core "problems" of variationist sociolinguistics. I argue that we can capitalize on the variability in pause realization (which we were mostly unable to account for earlier in this book by appealing to social factors), coupled with our knowledge that pause realizations have underlying relationships to the cognitive processes of speech production, to shed insight into the cognitive status of sociolinguistic variables.

I have examined the variable (ing) here and shown that speakers' realizations of (ing) bear a systematic relationship to the level of hesitancy during speech production. This hesitancy can be measured quantitatively through the slope lines generated in Henderson graphs and, as I have argued, can be linked to speakers' attention to their speech and used as a channel cue in speech production. The treatment here has been rather cursory and is intended primarily as a proposal. The evidence I have presented is, I believe, robust, but it only comes from one variable in one community. I hope that further studies can test this proposal and shed better light on the extent to which sociolinguistic variables, both (ing) and others, relate to temporal factors and ultimately to speakers' conscious (Chafe 1994) attention (Labov 1972) during speech production.

As Schilling-Estes (2002: 376) notes,

> Intra-speaker variation is pervasive, perhaps even universal, and we cannot hope to achieve a full understanding of the patterning of variation in language, or of language in general, if we do not understand its patterning within individuals' speech as well as across groups of speakers. Further, since intra-speaker variation lies at the intersection of the individual and the communal, a better understanding of its patterns will lend valuable insight into how the two spheres interrelate – that is, how individuals internalize broad-based community language patterns and how these patterns are shaped and re-shaped by individuals in everyday conversational interaction.

It is admitted that a focus on intra-speaker variation as something predicted or predictable from paralinguistic cues alone is overly reductionist and unrealistic. I have focused here on the relationship between measures of speech timing and the production of (ing) in terms of attention to speech, although I readily acknowledge a range of co-occurring factors, beyond speaker attention or intention or "focus of consciousness," are surely involved both in the production of the variable forms and the overall character of the talk – as measured through the Henderson graph metrics or likely through any single quantitative metric. It is clear that a range of other factors – environmental and speaker-internal distractions, nonlinguistic activities, etc. – play roles in speech and these are obviously missed by the strictly acoustically based timing measure implemented here. I take the fact that the Henderson graph-based slope measure succeeds as a predictive factor in the realization of (ing) as evidence of just how robust this phenomenon might be. All of the factors that it cannot account for (pauses in the acoustic signal due to noncognitive factors like nontalk action or interruptions that are not recoverable from the speech recording) should add noise to the analysis and distort the underlying patterns, yet the analysis shows a high level of systematicity despite the broad strokes that were used to generate the Henderson graph values.

Further, I am sensitive to the sort of critique of variationist approaches to style made by Coupland (2007), who argues that variationist work in general (and who might include the case study and Henderson graph method presented here) is overly correlational in nature. Coupland is right to argue that "the survey methods that Labov pioneered tend not to give priority to the local process through which [stylistic variation] happens. They orient much more to styles than to styling" (2007: 7).

He also makes a strong point that linguistic organization at the level of the sociolinguistic variable "isn't accessible to, or even directly relevant to, people engaged in speaking and listening, although it is the variationist's main concern. What matters to people is the meaning that language variation might add to their discursive practices – what people are trying to mean and what they hear others to be meaning" (2007: 8). Nonetheless, I believe this investigation has provided important insight into the psycholinguistic processes behind, and to a certain degree the reality of, attention to speech. It does seem to be the case that speakers (and likely listeners) craft individual variant forms as a part of their creation of referential and social meaning. These processes of sociolinguistic production have symptomatic traces in the speech stream, and, I believe, through accurate and finely focused attention to the sequential temporal patterns of talk sociolinguists can identify at least some of these symptoms, and through them better understand the psychological processes that underlie speech and broader social meaning-making practice.

Importantly, the Henderson graph method is not tied to the attention to speech model, or to any specific theoretical vantage beyond a belief that sequential temporal patterns of talk are bound up with language production at some level. It provides us a systematic means to assess intra-speaker variation in many of the terms recent papers have proposed, whether speaker design (cf. Schilling-Estes 2002, Coupland 2007) or "consciousness" in discourse (Chafe 1994) or, returning to its roots, something more centered in psychological studies (e.g. Arnold 2008). In sum, the Henderson graph method provides us a means to extend the systematic rigor associated with quantitative variationist methodologies to more current, but thus far most often qualitative, concerns.

Schilling-Estes (2002: 395) ends her survey of sociolinguistic approaches to style by noting that discovering the role of speaker awareness behind style is a pressing question and that maybe "attention to speech will once again come to the forefront." My proposal for the Henderson graph as a technique for studying aspects of sociolinguistic style is not necessarily to bring attention to speech to "the forefront" again but rather to highlight the psychological parallels between that approach to style and broader understandings of language production. The approach developed here is hoped to be one systematic step towards better understanding the roles of individual variation in sociocultural and linguistic processes overall.

More centrally related to the earlier parts of this book, we have seen that variability in pause – despite not being highly patterned itself by social factors – sheds light into the underlying systematicity of language variation. Pause realizations are tightly tied to cognitive activity during speech. We may not be able to reconstruct the patterns in pauses from a post hoc, corpus-based analysis of conversational speech, but we can leverage them to better uncover the processes at work during speech production, including those processes involved in the choice of (certain) sociolinguistic variables. We have examined only one sociolinguistic variable here, (ing), and clearly much more work is necessary to be able to draw stronger conclusions about the relationship between pausing and hesitancy – and the Henderson graph-based measures specifically – and the realization of sociolinguistic variation. My preliminary analyses of other sociolinguistic variables beyond (ing) (not discussed here) indicate that different variables have different relationships with the timing features measured through Henderson graphs. Ultimately, I believe, these different relationships will help us understand the different cognitive statuses of different variables. I hope you will agree that this is an exciting direction for variationist inquiry.[12]

9 Looking Back and Looking Further Forward

9.1 Taking stock

I began this book by noting that pauses and speech rate are ubiquitous features of speech, that they are characteristics of every utterance of every speaker and of every language. From the perspective of linguistics in general and sociolinguistics more specifically, it can be tempting to take these features for granted – to view them as either nonlinguistic (i.e. not of interest) or as simple and unidimensional. I hope the discussions and many studies of this book have illustrated that neither of these views are accurate. Silent pauses and rates of speech are complex phenomena with subtle patterns that shed light on social differentiation across and within speech communities and on the cognitive processes implicated in speech production.

Speech rate – as primarily viewed in terms of articulation rate, a measure of syllables per second excluding all pauses – was found to be highly linked to a wide range of social factors. Significant findings were found for speakers' regional background, ethnicity, sex, and age. In regional terms, we saw that speech rates appear to vary as much within a single region (e.g. between Central North Carolina and Southern North Carolina) as they do across regions (e.g. between Central North Carolina and Ohio), a finding which problematizes the common stereotype that Southerners talk slower than Northerners. We found other sociolinguistic effects as well, including those based on the speakers' interlocutors, in particular for the sex of the *interviewers* in the sociolinguistic interviews. We also saw evidence that some (but not all) speakers converge to their interlocutors' rates. Overall, the findings for rate of speech (qua articulation rate) indicate that its variability is highly systematic, even when viewed through a large-scale and somewhat coarse-grained corpus-based study.

Pause durations were found to be much less orderly in their relationships with social factors. Some social factors, mainly region and sex, were found to significantly influence pause durations, but, overall, the models that we could develop through our corpus-based methods at best poorly accounted for the pause variability. Interestingly, in terms of regional patterns, the pause findings do fall in line with the stereotype of slower-talking Southerners. Here we found that pauses were rather similar across all four subregions of North Carolina and that the Ohioans, the one Northern group in the data, had the shortest pauses (which corresponds to faster-sounding speech). More importantly, I argued – and I hope demonstrated – that all of this apparent noise when the pause data are viewed as dependent variables can be leveraged when these temporal factors are used as predictors of other phenomena. Pause variation is symptomatic of processes underlying speech production and can indeed be used as a channel cue, not only to attention to speech in the traditional, sociolinguistic sense but to speakers' cognitive processes at work as they speak, interact, and style themselves with others.

Recently, some sociolinguistic interest has moved away from the sort of macroscopic studies I have undertaken here.[1] In a groundbreaking paper commenting on the development of modern sociolinguistics, Penelope Eckert (2005, under review) classifies the study of sociolinguistic variation into three major categories, or "waves," and provides a nice overview of the major approaches to variationist sociolinguistics. The FIRST WAVE of sociolinguistics can be characterized by the study of broad correlational patterns between speakers' social features, like social class, sex, and ethnicity, and their use of variable language features. The SECOND WAVE of study involves ethnography and studying smaller groups of speakers to greater depth, focusing on more local patterns of language use, such as local affiliation or friendship networks. The THIRD WAVE of study is about practice and agency, rather than social structures. Instead of searching for categories which correlate with language use, research in the third wave focuses more closely on understanding styles and the construction and negotiation of identities rather than broad patterns of individual variable features. Eckert points out that these three waves are not necessarily chronologically ordered. Labov's (1963) groundbreaking first study – on Martha's Vineyard, with its deep ethnographic analysis of a small community – is a core example of the second wave, while his second (1966[2006]) foundational study – his large-scale survey of English in New York City – is squarely first wave. Yet, despite there not being a direct chronology that corresponds to the three waves, many scholars see current interest in sociolinguistics as moving increasingly towards third wave-like approaches (see also Coupland 2007).

One might readily classify the kind of approach I have undertaken here as first wave when viewed from this schema, or (regardless of schema) as hypercorrelational. I have indeed presented a broad-brush stroke over a large swath of data accumulated from a range of different community studies. However, I believe the kind of study I have undertaken here sheds crucial light on our understanding of language variation and change. As the study of sociolinguistics progresses, it is important that we balance approaches and use the most suitable frameworks for the relevant questions. For example, the third wave provides a deeper view of how meaning is made on the ground, and third wave studies are necessary as sociolinguistics develops better understandings of the meaning and meaning-making potential of language variation. The kinds of knowledge gained through third wave approaches are not possible through the kind of study I have undertaken. But, Eckert's critique is not, I believe, arguing that all variationist work should be third wave.

By taking advantage of the possibilities of large-scale corpus-based approaches we can draw new understandings about language and language variation and change from the massive accumulations of data that 50 years of productive sociolinguistic research have collected. All sociolinguistic, field-based studies, regardless of their original purpose, produce valuable naturalistic speech data. In aggregating these data we may lose the ability to apply the same insights gained through the initial fieldwork projects and ethnographies,[2] but we gain the ability to examine the robustness of socially based language variation across space and time. By viewing sociolinguistic datasets as corpora – or rather as data that can be aggregated and mobilized in corpus-like ways – we can, I argue, realize the benefits of third wave-like approaches and still gain the macro-level view that comes from the large-scale study of broader (albeit somewhat generic) social categories. (I again thank and acknowledge all of the researchers who have contributed to the data used in the studies I have presented – it is from these fieldworkers and their detailed and insightful fieldwork that sociolinguistically relevant corpus-based work becomes possible.)

Given the movement towards more ethnographically focused studies, some would surely critique the large-scale corpus-based approach I have taken as too near-sighted, and as sweeping too many things under the same rug. It is surely the case that my analyses have smoothed over many different factors in the data. For instance, clearly not all of the pauses in the dataset are of the same kind, and some are likely not even truly pauses during speech production – I have not been able to exclude or code separately silences that occur in the speech because a baby

cries in the next room or because a particularly rare and beautiful bird lands outside a living room window. These facts of the interview data (unless explicitly marked in the transcripts) are beyond the ability of the analysis to account for. Yet, these sorts of issues, from a methodological standpoint, are just additional sources of noise and should make the statistical analyses perform less well. The fact that the analyses do as well as they do is, I believe, a testament to the reality of the findings.

This is especially true of the Henderson graph analysis of Chapter 8. There I used a somewhat coarse technique for computing the slope lines (i.e. the pause-to-talk metrics) and a particularly conservative measure of slope (the best-fit linear models used to calculate the slope smooth over the more extreme variability in the pause/talk alternations). Both of these facts should disrupt our ability to find meaningful patterns in the data – the former for adding unsystematic noise in the data and the latter for limiting the range of the obtained slope values. Yet, systematic findings emerge. The most powerful aspect of the Henderson graph method, I believe, is its ability to let us compare different sociolinguistic variables and different speakers and to test new hypotheses about language variation and cognition. It may be the case that this method is not successful or that others are found that shed better light on these relationships. Either way, I put forth the idea here as a proposal, as a suggestion for new directions for sociolinguistic inquiry. Whatever we learn, I believe, will move us forward on the path to understanding the social *and* psycholinguistic bases of language variation.

I expect that the method can tell us even more when coupled closely with third wave-like ethnographic work (even if the local insight is simply used to better compute the slope lines and better determine whether some passages should be excluded from analysis due to outside distractions or other problems – though it seems to me that very much more is possible). By getting closer to meaning-making "on the ground," third wave approaches could help us better connect the dots between language variation, identity, and cognition. Ultimately, a combination of methods – ethnographic, corpus analytical, and experimental – will be necessary for uncovering the fuller nature of linguistic variation.

I end this book by acknowledging that there are many more questions than answers and hoping that I have adequately explicated some of these questions and shared my excitement for their pursuit. At the very least, it is my hope that the studies in this book have pointed to the usefulness of the large-scale, sociophonetic analysis of less studied speech features, and the benefit of treating our collected sociolinguistic data in corpus-based ways.

Appendix I: Guide to the Website

In order to make this book more accessible and useful to its readers, I have placed many resources related to this project on the book's website. These include downloadable tools, like the syllable-counting algorithm ported as a function in the R language, downloadable datasets from the book, including Tables 4a and 5a (information about each of the speakers examined in Chapters 4 and 5), and a web-based Henderson graphing tool for generating graphs and metrics like those used in Chapter 8. Readers are also referred to the website for future addenda and additional information.

Website address: http://ncslaap.lib.ncsu.edu/speechrateandpause/

Appendix II: Correspondences between log-millisecond (log-ms) and millisecond (ms) pause durations

Throughout the pause duration analyses in this book I often refer to pause times in log-converted measures. The table here provides some simple correspondences between log-ms and ms pause measures. All logs in this book are natural logs (i.e. to the base *e*).

log-ms	**ms**
3.912	50
4.605	100
5.298	200
5.704	300
5.991	400
6.215	500
6.397	600
6.551	700
6.685	800
6.802	900
6.908	1000
7.090	1200
7.244	1400
7.378	1600
7.496	1800
7.601	2000
8.006	3000
8.294	4000
8.517	5000

Notes

1 Looking Forward

1. SLAAP's website is http://ncslaap.lib.ncsu.edu/. SLAAP is discussed further in Chapter 3.
2. See Clark and Fox Tree (2002) and the papers that followed and debated it for an interesting line of research on the semantics of *uh* and *um*.
3. But not always. See Poplack (1989) and projects like, for example, Tagliamonte (2008) which collect data and build sociolinguistic corpora explicitly for long-term use.
4. /r/-ful, or rhotic, meaning pronouncing /r/s in words like *car* and *bird*. New York City is famously an /r/-less dialect area (cf. Labov 1966[2006]).
5. COGNITIVE SOCIOLINGUISTICS is a term that has recently had some traction and has some obvious loose associations with the case I am making here (cf. Kristiansen and Dirven 2008, Geeraerts, Kristiansen, and Peirsman 2010). Yet, this term has primarily been connected with the (capital C, capital L) Cognitive Linguistics movement, which is more theory-specific than I wish to be in these pages. Cognitive Linguistics (CL) – and its socio-branch – have a particular usage-based theoretic stance and a foundation in pragmatics and semantics. While interesting, it is not as connected or connectable to the growing line of research in sociophonetics that intersects traditional research in sociolinguistics and psycholinguistics. See Thomas (2011a: 300–1, 2011b) for further discussions.
6. See Kendall (2011) for a consideration of the relationship between sociolinguistics and corpus linguistics.

2 What We Know about Speech Rate and Pause

1. Or, if of interest, as PARALINGUISTIC. As I will describe in Chapter 8, Labov famously considered them potential paralinguistic channel cues to a speaker's style qua attention to speech.
2. To DRAWL: "to prolong or lengthen out the sounds of speech in an indolent or affected manner; to speak slowly, by affectedly prolonging the words" (*Oxford English Dictionary*). Surprisingly, few researchers have actually investigated what features specifically contribute to the "drawl" (Sledd 1966, Feagin 1987, Wetzell 2000).
3. I thank John Singler for reminding me about this excerpt.
4. Wells's full quote reads: "It is perhaps universally true that rural accents tend to be slower in tempo, reflecting the unhurried life of the countryside: compare a New York and a hillbilly accent, or Cockney and Wiltshire. Urban accents tend to be not only faster, but also more up-to-date in terms of sound change in current progress" (1982: 11).
5. And the full quote here: "Accents also vary in pace: that is, in the general rate of speech, measurable as the mean number of syllables uttered per second.

There exists a general tendency for urban speech to be faster than rural speech, as can be seen by comparing the accent of London with that of Wiltshire, that of New York City with that of Texas, or that of Melbourne with that of rural Australia. It must be confessed, however, that to state the existence of this tendency is to make an impressionistic claim rather than to report a substantiated fact. In any case, it is only a general tendency. Particular urban or rural areas do not necessarily conform to it. Particular speakers from a given geographical background [do] not necessarily conform, except in a general statistical way, to the norm for their locality. And any individual varies his pace of utterance in accordance with situational factors or personal whim" (Wells 1982: 87).

6. However, Kowal and O'Connell (1980) report conflicting results with respect to pause location – that pauses are more frequently aligned with function words and not content words – and argue that Maclay and Osgood's (1959) methodology was problematic.
7. And, as will be discussed in the next chapter, pause plays a primary role in the determination of what constitutes an "utterance" (a transcript line) within SLAAP's transcription system and ultimately the scope of each articulation rate measurement.
8. Before ending this section, I should quickly acknowledge that I have not discussed FILLED PAUSES, like *eh*, *uh*, and *um*, in this section and that they are for the most part excluded from the considerations of this book. Filled pauses are quite interesting and many studies have shed light on their patterns and meanings (cf. Rochester 1973, Clark and Fox Tree 2002, Campione and Véronis 2005, etc.). However, examining filled pause phenomena is outside the scope of the present project.
9. Ray and Zahn (1990) present a "preliminary" (in their own terms) look at a study of regional patterns of speech rate in the US. As noted, they do not find significant differences by region or by gender, but it is also worth noting that their methods were perhaps not sophisticated enough to find these differences. Their study examined only samples of a minute or two of speech from their subjects and calculated rates in terms of words per minute.
10. Jacewicz et al. (2010) also consider whether their differences from Quené (2008) are a result of using different units for the analysis. To investigate this, they reevaluate their data in terms of average syllable duration (ASD), an alternate to a measure of syllables per second, and find that "no matter which of the two approaches is taken in modeling overall speech tempo or phrase length ... the obtained results will be basically the same" (Jacewicz et al. 2010: 846). In testing this, they provide a nice confirmation that the decision of which measurement unit to use alone does not influence the main outcome of the analysis. Nonetheless, as I will discuss in §4.4.1, the use of different measurement units does involve different levels of precision and so may still have some influence on the comparability of results.
11. And, to what extent does statistically significant variation matter if it is below the level of speakers' perception (Labov 1994)?

3 New Tools and Speech Databases

1. The NCLLP is a sociolinguistic research initiative directed by Walt Wolfram at North Carolina State University with one of the largest audio collections

of sociolinguistic data on Southern American English in the world. The growing collection contains several thousand interviews conducted from the late 1960s up to the time of this writing, most on analog cassette tape, but some in formats ranging from reel-to-reel tape to digital video. (For more information about the NCLLP visit the project's website at http://www.ncsu.edu/linguistics/ncllp/.) The NCLLP's large and growing collection of interviews is an important resource for linguists in general and for other scholars interested in the American South. As a part of the SLAAP initiative, almost all of the legacy sociolinguistic interviews have been digitized and catalogued and new recordings are added to SLAAP as they are collected.

2. SLAAP currently only houses audio materials, although the software has been designed to be extensible in this regard, and it is hoped that video recordings from the NCLLP collection and elsewhere can be incorporated into the archive in the near future.
3. SLAAP's method for calculating speech rate is discussed in §4.4.1.
4. See also Barbiers, Cornips, and Kunst (2007) on using Praat for time-aligned transcription. MacWhinney (2007) discusses some other tools that can be used for this purpose, such as Transcriber and CLAN. Kendall (2006–2007, 2009), MacWhinney (2007), and others (e.g. Edwards 2001) provide general discussions on the benefits of time-aligned transcripts.
5. Also see Preston (1982, 2000) on the importance of the choice of orthography in transcription.
6. "Talk-Time" is the measure of how much actual talk (phonation) is made by a given speaker. This measure does not include silent pauses. "Turn-Time" is the measure of how long a given speaker "holds the floor," which includes intra-turn silent pauses. These figures are generated dynamically by the software. The amount of inter-speaker pauses and silence – i.e. gap length – can be reconstructed by subtracting the total "Turn-Time" from the total length of a transcript. In other words, just over 5 percent of the transcribed interview consists of inter-speaker pauses, while about 21 percent of the transcribed interview (865.24 sec – 678.01 sec / 911.00 sec) consists of intra-speaker pauses. So, overall, just over a quarter of the transcribed interview consists of silence.

4 Methods and a First Look at Speech Rate and Pause

1. I am grateful to Valerie for her collaboration on our larger projects involving these data and for her comments on an early draft of this chapter.
2. I should also acknowledge that developments in the statistical analysis of linguistic data continue to occur at a rapid rate and researchers do not always agree on the absolute best practices. The statistical methods I use are based on widely discussed and published approaches (again, primarily those articulated in Baayen 2008). It may be the case that readers find fault with some of my decisions – this seems to me inevitable in a time when methods continue to be developed and negotiated. My hope is that by being methodologically consistent throughout the projects presented in this book, and by being explicit about the approaches I have taken throughout, I have made this work more transparent and the findings more interpretable, even if some readers disagree with choices I have made.

3. Other statistical packages are available that also support powerful regression but I focus on R as it is in wide and growing use at the time of this writing, and its free status, open-source development model, and global popularity across disciplines, indicate a promising future for the software (cf. Baayen 2008, K. Johnson 2008, Gries 2009).
4. However, in Chapter 8 I use a slope-based measure of hesitancy to look more deeply at speech timing as an independent variable for the analysis of sociolinguistic variables.
5. To risk getting ahead of myself, these utterance length figures are much higher than what we find in most conversational speech. The data examined in the next chapter, just over 30,000 measurements from conversational sociolinguistic interviews, yield a mean syllables per utterance count of 6.96 and a median value of 6.0.
6. Boxplots, or box-and-whisker plots, will be used throughout this book and are worth a brief explanation for readers unfamiliar with them. This plotting technique provides a quick, and I believe easily interpretable, summary of numerical data. The dark band in the center of each box shows the median value of the distribution. The "box" surrounds the first and third quartiles of the data and the "whiskers" extend to include the furthest data points that are no more than 1.5 times the interquartile range (i.e. the height of the "box"). Outlier points, points outside of 1.5 interquartiles from the box, are shown as individual data points, as in the left-hand panel in Figure 4.6. Benjamini (1988) provides a nice, accessible discussion of the generation and interpretation of boxplots. The boxplots I generate throughout this book are created using the standard `boxplot()` function in R.
7. Based on the discussions in Chapter 2 (e.g. Tannen's 1984[2005] CONVERSATIONAL STYLES), we might expect a range of additional factors to influence within-region speech rates, such as cultural orientation, religious affiliation, and so on. However, recall that all of the reading passage data examined here come from European Americans of roughly the same age. We are unable to look in depth at socially based variation within the regions in these data. The only within-region social factor available for these data, speaker sex, does not arise as significant in the analysis.
8. Following typical practice, we interpret p values of less than 0.05 as statistically significant. At certain points later in this book, I occasionally attend to p values equaling (even when rounded down to) 0.05 and treat them as "on the cusp of significance." I only include these truly marginal p values in, for instance, the "best" statistical model of a feature when the inclusion of the factor is found to significantly improve the model through other assessments; that is, the marginal p value alone does not warrant its inclusion.
9. A linear regression is used here because the independent variable, syllables per utterance, is a continuous predictor.
10. This figure is generated using Harald Baayen's (2008) R function `xylowess.fnc()`. The dots represent individual measurements, while the solid line in each plot panel depicts a lowess smoothed best-fit line for the data (Cleveland 1981).
11. The confidence intervals in the model plots are generated through the Monte Carlo estimation of p values.

12. Often when modeling continuous predictors like the age of speakers, or, as here, NUMSYLS and the START time of the utterances, the factors are CENTERED so that their mean values are set to zero. However, I have not centered factors in most of the analyses in this book and a brief word about CENTERING is in order. The regression models assume a default value of zero for continuous predictors and the models' intercepts – that is, their baseline prediction for the dependent variable when all factors have their default values – are estimated based on these assumed zeros. Yet, zero is often not a realistic or meaningful default value for a continuous predictor. For instance, no utterances have zero number of syllables and no speakers have an age of zero. Centering allows the model to obtain a more meaningful intercept by using the continuous predictors' mean value as the zero point. Centering does not, however, otherwise impact the model results beyond influencing the intercept, and it can have the undesirable effect of making the values for the centered predictors harder to interpret. In this book, we are interested in interpreting the effects of the predictors more than the models' intercepts, and, as such, I leave these factors uncentered.

5 Speech Rate and Pause in Conversational Interviews

1. This interviewer, whom I label "C" later in the book, is not included in the analyses of speech rate and pause in this chapter. We will examine her speech, however, in Chapter 7, where I take advantage of this large amount of transcribed talk to better understand intra-speaker variation and accommodation effects.
2. It could be argued based on North Carolina's sociohistorical development and dialect distribution that Warren County, in the north of the state, should be considered separately from the other locations deemed Central NC. This would be a reasonable argument. For the time being, I have defined the Central NC region in negative terms – not the western mountains, not the Atlantic coast, not the uniquely triethnic Robeson County in the south of the state. I have also chosen to keep the Warren County speakers within the Central NC category to better balance the distribution of speakers. Future work will need to assess whether the Warren County speakers can be differentiated from the Central NC speakers to their south.
3. In the later statistical analysis of the data, age as an independent variable slightly outperforms year of birth, so I stick to this view through the presentations and discussions of the data. The finding that age is a better predictor than year of birth supports a notion that speech rate and pause are AGE-GRADED features and not undergoing change over time.
4. In the per-speaker model later in this chapter, we will test the effect of each speaker's median pause duration on their articulation rate (and vice versa), but this is not done here. At the level of the individual utterance, there is no coordinated measure of pause duration to associate with utterance. Some utterances follow and/or precede a speaker's intra-turn pause, but other utterances are bounded by talk by other speakers, or by nontalk action.
5. The plot is generated by the `summary()` function in the Design library in R (Harrell 2009).

6. In this and the following tables square brackets are used to differentiate nonsignificant *p* values from significant *p* values. Brackets in italics are used to highlight *p* values on the cusp of significance (≈ 0.05). As in Table 4.2, nonlinear components are indicated using ' marks. See Chapter 4 for a discussion of how these nonlinearities in the data are modeled using RESTRICTED CUBIC SPLINES (Baayen 2008: 174–81).
7. In Table 4.2, earlier, I explicitly noted the contrast for each dummy coded categorical factor (e.g. "North (not West)"). I have not done this here for sake of space. For REGION, the baseline factor is central NC, for SEX it is female, and for ETHNICITY it is African American. The rows for these factors are comparisons against these baselines.
8. The number of "knots" used by the spline function to fit nonlinear continuous factors is determined through the same techniques used for selecting the factors to include in the best statistical model, as was discussed in §4.2; models are built which vary in only the number of knots included in a continuous factor fitted with a spline and likelihood ratio tests are used on pairs of models differing only by one knot to determine whether the additional knot increases the fit of the model enough to warrant the cost of the additional parameter. Simply put, the number of knots determines how many bends occur in the otherwise linear prediction (Baayen 2008: 174–81; see further Harrell 2001).
9. To generate this validation measure, I iterated 200 times over a randomly sampled subset (with replacement) of the 29,600 measurements of the trimmed dataset and generated an R^2 value for each iteration. I then estimated the overall R^2 by taking the mean of these 200 separate R^2 measures.
10. This value is different than the mean value of age for the articulation rate data because each speaker contributes different amounts of pauses and utterances to the dataset and the mean age is calculated based on the age associated with each utterance (for speech rate) and pause (for pause duration), and thus speakers' weights towards the mean age are dependent on their *N*s.
11. Articulation rate is measured excluding pauses, so articulation rate and pause duration values are not correlated, at least not essentially. Any relationship between the two will be indicative of larger, linguistic or physiological constraints.
12. The stylistic differences in the two sets of figures – the model results for the mixed-effect model, Figure 5.3, and the model results for the fixed-effect regression, Figure 5.8 – are a result of the different plotting functions that must be used in R to plot the results from the two types of regressions. I apologize that the formats of the two sets of figures are not more similar.
13. For the fixed-effect models, the Design library's `validate()` function performs the bootstrap validation technique (Harrell 2009). As for my manually calculated bootstrap validation performed for the model of Table 5.2, this validation is tested over 200 iterations.
14. As before, since articulation rate is measured exclusive of pause articulation rate and pause duration values are not correlated, at least not essentially.
15. We can imagine many different ways to incorporate a pause frequency count into the analysis. However, the count must be normalized in some way, since we have different amounts of data from different speakers. I have normalized this to a measure per 100 words, but we could also examine, for instance,

measures normalized per X syllables or per X seconds of phonation. I have not tested every possible frequency measure and, for this analysis, pursue only the pauses per X words measure. (In all cases, the number of units, X, does not matter, e.g. the difference between normalizing by 100 words and 1000 words only changes the place of the decimal.)

16. Speakers' pause durations did not, however, predict their articulation rates in the last section. It is perhaps the case that the minor relationship between pause and articulation rate is masked in the articulation rate model by the better accounting for the variance that the other factors provide.
17. Overfitting means that a statistical model describes the data it is fit to too specifically and does not generalize to unseen data. The goal in regression modeling – especially as I have employed it – is in making predictions about a larger population from a given sample. If the model overfits the data, it does not accurately describe the larger population. The bootstrap validation used in §5.3 tests for overfitting. The adjusted R^2 values presented in §5.4 penalize for overfitting, which is why they are noticeably lower than the raw R^2 values for Tables 5.5 and 5.6.

6 Closer Looks at Speech Rate and Pause Variation: Methods and Findings

1. This is done using the `sample()` function in R, which performs a standard randomization algorithm. For each iteration this function is used to randomly select from the original data a sample ten measurements smaller than the previous iteration. Another way to do this would be to randomly remove ten samples from the same dataset each iteration and I did this as well, obtaining very similar results. I have not shown this alternative version.
2. I am focusing here on determining the best minimum pause threshold value for the analysis of the pause duration data only. Since pauses are used to separate the phonetic utterances in the articulation rate data, the question of the minimum pause duration is an important – though different – question for that analysis as well. There, I have used the smallest possible criteria for pause durations – 60 ms – in the belief that removing as much silence from the analyzed talk gives the most precise measure of *articulation rate* (the pause-exclusive measure). Robb et al. (2004) removed pauses about this short (50 ms, in fact) in their analysis of articulation rate. Many studies of articulation rate do not report what length of silences they count as pauses to be excluded from their "speech" data for articulation rate calculations.
3. Of course, this approach to model comparison is only useable for comparing the same model (i.e. models with the same parameters) on very similar datasets. Otherwise, we would be comparing apples with oranges, so to speak.
4. Intonational Phrases (IPs) are defined here following the main ToBI literature (Beckman and Hirschberg 1994, Beckman, Hirschberg, and Shattuck-Hufnagel 2005; see Thomas 2011a: 200–18 for a sociophonetically oriented introduction). An IP is the highest level of phrase in intonation and typically characterized by a clearly identified edge tone, a resetting of pitch to mark the start of each IP, and sometimes, but not always, by adjacency to a pause.

5. I am indebted to Erik for his participation on this part of the project. This section – and in fact much of this book – would not have been possible without his collaboration and insightful comments. The data used in this section were coded by hand and are partly used as an accuracy check for the more automated measured of the other sections. However, it must be acknowledged that "the impressionistic nature of intonational transcription and the volume of speech transcribed this way here make some degree of error inevitable" (Thomas personal communication).
6. One utterance in these data was coded as having a final foot of five syllables, but this was excluded from the analysis since such a long final foot only occurred once.
7. These are the same speakers who were described as European American and Latino/a, respectively, earlier. The issue of naming different ethnic groups is often complex. In the South Texas context, Anglo and Hispanic are more preferred terms and, following my earlier work with Erik Thomas (Kendall and Thomas 2010) on these particular data, I use these terms here.
8. I thank Brendan Brown, at the University of Oregon, for help with data coding for this part of the study.

7 Closer Looks at Speech Rate and Pause Variation: Interlocutors and Accommodation

1. Readers unfamiliar with sociolinguistic interviews may wish to read Labov (1972) or a more recent sociolinguistic research guide (like Milroy and Gordon 2003) for an overview of this data collection technique. In short, sociolinguistic interviews seek to maximize the naturalness of the situation (but also often elicit a range of speech tasks, including reading passages and word lists; though it should be noted that beyond Chapter 4 the data used in this book come only from the conversational portions – the "interview" part – of the recordings). Often having more participants present than just a single interviewer and interviewee increases the interviewee's comfort level. Sociolinguists are also often quite happy to have an "interloper" join in the conversation as this can further promote more natural conversational speech.
2. I have used female as the baseline factor for the IntvrSex factor to parallel the treatment of interviewee Sex here and throughout this work. A comparable model built with male as the baseline for IntvrSex yields the same significant difference for females (as it should; estimate = –0.20; $p = 0.033$) but shows that "mixed" interviewers are also not significantly different than male interviewers ($p = 0.563$). Thus, interviewees' articulation rates when they have male *and* female interviewers (together) fall between their rates when they are interviewed by just males or just females.
3. I am extremely grateful to Danica for her letting me use her data and draw on some of her analysis. Of course any errors in the analysis presented here are my own.
4. I am grateful to Carissa Froyum Roise for sharing these data with Christine Mallinson and me. As I will acknowledge again in Chapter 8, I am extremely thankful to Christine for all of her work on these data.

5. (–) and (+) are used in these tables to indicate one standard deviation below or above the group's mean, respectively.
6. These means are computed from the median values from each interview. They are very similar, however, to C's overall rates: her overall median pause duration is 400 ms (calculated from the individual 3181 pause measurements) and her overall median speech rate is 4.87 σ/sec (from the 4725 speech rate measurements).
7. As I briefly explore in Kendall (2010b), some of the possible accommodation effects may be found to relate to speakers' *perceptions* of their interlocutors more than they do to actual aspects of their interlocutors' speech – as we have focused on here. Again, following up on this is left for work elsewhere.

8 The Influence of Speech Rate and Pause on Sociolinguistic Variables

1. However, many researchers found the "danger of death" prompt itself to be unsuccessful in the field and sometimes even disastrous. See Butters (2000) for a discussion.
2. Labov's was not the only early work undertaken on "style" from sociolinguistic perspectives (see e.g. Giles 1973, Hymes 1974, Giles and Powesland 1975), but the other contemporary approaches do not appear to have lent themselves to the burgeoning field of quantitative study pioneered by Labov and did not achieve the same uptake as Labov's attention to speech model in the early quantitative work.
3. The issue of comparability across interviews and interview segments is nicely taken up, from a different perspective, in recent work by Frans Gregersen and colleagues (cf. Gregersen, Beck Nielsen, and Thøgersen 2009, Gregersen and Barner-Rasmussen 2011).
4. Some of this section reiterates points (and citations) from Chapter 2. I risk being overly repetitive to keep this chapter somewhat self-contained.
5. Guy's recent work (personal communication) has returned to this question, and found some evidence in support of the relationship between rate of speech and deletion.
6. The notion of COGNITIVE RHYTHM turned out to be quite controversial and has been debated at some length in the years following Henderson et al.'s (1966) publication (e.g. Jaffe, Breskin, and Gerstman 1972, Henderson 1974, Butterworth and Goldman-Eisler 1979, Power 1983, Beattie 1984, Power 1984, Kowal and O'Connell 1985, and so forth). For our purposes, I am less interested in whether the alternation between planning and execution phases claimed by Henderson et al. (1966) is a "real" phenomenon and can be accurately discerned through a Henderson graph-like approach. Rather, I am interested in the empirical power of the graphing technique and the more general – and less controversial – idea that a measure of pause-to-talk ratio can act as a proxy to ongoing cognitive processes during speech production.
7. In the actual statistical analyses discussed shortly, I log-transform SLOPE (LSLOPE) to account for the fact that it limits zero, and like pause durations, has a log-normal distribution. I also center the log-transformed LSLOPE (CLSLOPE) for each speaker individually. This CLSLOPE variable then becomes a relative measure of speakers' slopes, somewhat similar to SLOPECOMP, rather than an absolute measure.

8. With the exception of SLOPECOMP and ARTRATECOMP, all of the temporal sequencing measures in Table 8.1 are conceptualized as continuous variables. SLOPECOMP and ARTRATECOMP are calculated based on whether individual values are above ("high") or below ("low") one-half a standard deviation from the speaker's mean, with values within a half standard deviation considered "norm(al)." This is of course just one possible way to categorize the different values. Once the Henderson graphs are built, the possibilities for deriving potential independent variables are expansive. The challenge lies in determining which of the many potential variables are meaningful and "real" and which are spurious artifacts of the procedure that do not reflect or usefully capture characteristics of the actual talk.
9. I am grateful to Christine Mallinson for all of her collaborative work on these data, as well as to Kaye Wise Whitehead who also helped us with some of the data. All of the variable and discourse data I examine here, beyond the Henderson graph-derived measures, are a result of Christine's and my many collaborations and would not have been possible without her, her hard work, and her insights. Any errors here are, however, entirely my own fault.
10. Variable (ing) is often described as VELAR NASAL FRONTING, a term which implies an active, articulatory fronting of the nasal segment. It is also known as G-DROPPING, although the alternation does not actually involve the "dropping" of a segment. The extent to which the alternation is a phonological variable (e.g. involving articulatory fronting), a morphological variable (e.g. involving an alternation between allomorphs), or a lexical alternation is beyond the interest of this consideration. I sometimes will describe the alveolar form, *-in'*, as fronted or as fronting and the velar form, *-ing*, as backed but I mean this in purely descriptive terms. The vowel quality can also vary depending on variety, but we follow the common practice of impressionistically coding the variant based on its consonant.
11. This very simple linguistic model of the (ing) data does not attempt to assess all of the possible independent linguistic factors behind (ing) realization. The analysis here is meant to assess and illustrate the uses of Henderson graph-based metrics and as such I have limited the depth and breadth of analysis into other linguistic factors. See Tagliamonte (2004) for a fuller review of factors behind (ing) realization. In Kendall (2010b), I also examine these (ing) data. However, those data, and the models presented in that work, are drawn from a larger dataset so yield slightly different statistical results. The data examined here are pruned to only include instances of (ing) that fall within measurable Henderson segments.
12. I have made a public tool for Henderson graphing available from the book's website and it is my hope that others will test the method on their data and on new variables. I also hope that future users will discover better, more accurate ways to calculate the Henderson slopes and the graphs' other derived measures. Surely, better tools than I have thus far developed can be built and can better advance this enterprise.

9 Looking Back and Looking Further Forward

1. This consideration of the three waves of sociolinguistic study draws from my paper "Corpora from a sociolinguistic perspective" (Kendall 2011), where I consider the simultaneous movements towards and away from corpora by sociolinguists.
2. When drawing data from, for instance, two disparate communities with different social histories and hierarchies, and different original research designs and foci, it must be accepted that not all aspects of the data will be comparable or should be compared.

References

Alvarez, Louis and Andrew Kolker. 1988. *American Tongues*. New York: The Center for New American Media. Video.

Apple, William, Lynn A. Streeter, and Robert M. Krauss. 1979. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology* 37: 715–27.

Armstrong, Meghan, Cynthia Clopper, and Rajka Smiljanic. 2008. Pause distribution in three American English dialects: Southern, Midland and Mid-Atlantic. Paper presented at New Ways of Analyzing Variation (NWAV) 37: Houston, TX.

Arnold, Jennifer E. 2008. Reference production: Production-internal and address-oriented processes. *Language and Cognitive Processes* 23(4): 495–527.

Arnold, Jennifer E., Maria Fagnano, and Michael K. Tanenhaus. 2003. Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research* 32(1): 25–36.

Arnold, Jennifer E., Carla L. Hudson Kam, and Michael K. Tanenhaus. 2007. If you say *thee uh* you are describing something hard: the on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(5): 914–30.

Austin, J. L. 1962[1975]. *How to Do Things with Words*, 2nd edn. Cambridge, MA: Harvard University Press.

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: a Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bailey, Guy. 2002. Real and apparent time. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Malden, MA/Oxford: Blackwell, 312–32.

Baker, Rachel and Valerie Hazan. 2010. LUCID: a corpus of spontaneous and read clear speech in British English. *Proceedings of DiSS–LPSS Joint Workshop 2010*. Tokyo, Japan, 3–6.

Barbiers, Sjef, Leonie Cornips, and Jan Pieter Kunst. 2007. The Syntactic Atlas of the Dutch Dialects (SAND): a corpus of elicited speech and text as an online dynamic atlas. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, 54–90.

Bartošek, Jaroslav. 1974. *Jazyk Rozhlasové Sportovní Publicistiky*. Prague.

Bassett, Mary, Daniel O'Connell, and William Monahan. 1977. Pausological aspects of children's narratives. *Bulletin of the Psychonomic Society* 9: 166–8.

Bates, Douglas and Martin Maechler. 2010. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-37.

Beal, Joan, Karen Corrigan, and Hermann Moisl (eds). 2007a. *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan.

Beal, Joan, Karen Corrigan, and Hermann Moisl (eds). 2007b. *Creating and Digitizing Language Corpora*. Vol. 2: *Diachronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan.

Beal, Joan, Karen Corrigan, and Hermann Moisl. 2007c. Taming digital voices and texts: Models and methods for handling unconventional synchronic corpora. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, 1–16.
Beattie, Geoffrey. 1984. Are there cognitive rhythms in speech? – A reply to Power (1983). *Language and Speech* 27: 193–5.
Beckman, Mary E. and Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston and Mary E. Beckman (eds), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge, UK: Cambridge University Press, 152–78.
Beckman, Mary E. and Julia Hirschberg. 1994. The ToBI Annotation Conventions. http://www.ling.ohio-state.edu/~tobi/ame_tobi/annotation_conventions.html
Beckman, Mary E., Julia Hirschberg, and Stephanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), *Prosodic Typology: the Phonology of Intonation and Phrasing*. New York/Oxford: Oxford University Press, 9–54.
Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea. 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113(2): 1001–24.
Bell, Allan. 1984. Language style as audience design. *Language in Society* 13: 145–204.
Bell, Allan. 2001. Back in style: Revisiting audience design. In Penelope Eckert and John R. Rickford (eds), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, 139–69.
Benjamini, Yoav. 1988. Opening the box of a boxplot. *The American Statistician* 42(4): 257–62.
Block, Susan and Deanne Killen. 1996. Speech rates of Australian English-speaking children and adults. *Australian Journal of Human Communication Disorders* 24(1): 39–44.
Boersma, Paul and David Weenink. 2010. Praat: Doing phonetics by computer. http://www.praat.org/
Boomer, D. S. 1970. Review of Frieda Goldman-Eisler, *Psycholinguistics: Experiments in Spontaneous Speech*. *Lingua* 25: 152–64.
Bourdieu, Pierre and Luc Boltanski. 1975. Le fêtichisme de la langue. *Actes de la Recherche en Sciences Sociales* 4: 2–32.
Brown, Bruce. 1980. Effects of speech rate on personality attributions and competency evaluations. In Howard Giles, William P. Robinson, and Philip M. Smith (eds), *Language: Social Psychological Perspectives*. New York: Pergamon, 293–300.
Butcher, Andrew. 1981. *Aspects of the Speech Pause: Phonetic Correlates and Communicative Function*. Kiel, Germany: University of Kiel.
Butters, Ronald R. 2000. Conversational anomalies in eliciting danger-of-death narratives. *Southern Journal of Linguistics* 24(1): 69–81.
Butterworth, Brian and Frieda Goldman-Eisler. 1979. Recent studies on cognitive rhythm. In Aron W. Siegman and Stanley Feldstein (eds), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 211–24.

Byrd, Dani. 1994. Relations of sex and dialect to reduction. *Speech Communication* 15(1–2): 39–54.
Campbell-Kibler, Kathryn. 2005. Listener perceptions of sociolinguistic variables: the case of (ING). PhD dissertation. Stanford University.
Campbell-Kibler, Kathryn. 2007. Accent, (ING) and the social logic of listener perceptions. *American Speech* 82(1): 32–64.
Campbell-Kibler, Kathryn. 2010. Perception in sociolinguistics. *Language and Linguistics Compass* 4(6): 377–89.
Campione, Estelle and Jean Véronis. 2002. A large-scale multilingual study of silent pause duration. In Bernard Bel and Isabelle Marlien (eds), *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence: Laboratoire Parole et Langage, 199–202.
Campione, Estelle and Jean Véronis. 2005. Pauses and hesitations in French spontaneous speech. *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop*. Aix-en-Provence, France, 43–6.
Cappella, Joseph. 1985. Controlling the floor in conversation. In Aron W. Siegman and Stanley Feldstein (eds), *Multichannel Integrations of Nonverbal Behavior*. Hillsdale, NJ: Lawrence Erlbaum, 69–103.
Carpenter, Jeannine. 2005. The invisible community of the Lost Colony: African American English on Roanoke Island. *American Speech* 80: 227–56.
Cedergren, Henrietta and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2): 333–55.
Chafe, Wallace. 1980a. The deployment of consciousness in the production of a narrative. In Wallace Chafe (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex, 9–50.
Chafe, Wallace. 1980b. Some reasons for hesitating. In Hans Dechert and Manfred Raupach (eds), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton, 169–80.
Chafe, Wallace. 1985. Some reasons for hesitating. In Deborah Tannen and Muriel Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, 77–89.
Chafe, Wallace. 1993. Prosodic and functional units of Language. In Jane Edwards and Martin D. Lampert (eds), *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum, 33–43.
Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: the Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago.
Clark, Eve. 2009. *First Language Acquisition*, 2nd edn. Cambridge: Cambridge University Press.
Clark, Herbert and Jean Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition* 84: 73–111.
Cleveland, William. 1981. LOWESS: a program for smoothing scatterplots by robust locally weighted regression. *The American Statistician* 35: 54.
Cleveland, William. 1993. *Visualizing Data*. Summit, NJ: Hobart Press.
Clopper, Cynthia and Rajka Smiljanic. 2011. Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics* 39: 237–45.
Coupland, Nikolas. 2007. *Style: Language Variation and Identity*. Cambridge: Cambridge University Press.
Crown, Cynthia and Stanley Feldstein. 1981. Conversational congruence: Measurement and meaning. Paper presented at the Eastern Psychological Association: New York.

Crown, Cynthia and Stanley Feldstein. 1985. Psychological correlates of silence and sound in conversational interaction. In Deborah Tannen and Muriel Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, 31–54.
Cullinan, Danica. 2007. Intraspeaker variation in the sociolinguistic interview. Unpublished manuscript. Raleigh, NC: North Carolina State University.
Dankovičová, Jana. 2001. *The Linguistic Basis of Articulation Rate Variation in Czech*. Frankfurt am Main: T. Hector.
De Jong, Nivja and Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41 (2): 385–90.
Deese, James. 1984. *Thought into Speech: the Psychology of a Language*. Englewood Cliffs, NJ: Prentice-Hall.
Di Paolo, Marianna and Malcah Yaeger-Dror (eds). 2011. *Sociophonetics: a Student's Guide*. London: Routledge.
Drager, Katie. 2010. Sociophonetic variation in speech perception. *Language and Linguistics Compass* 4(7): 473–80.
Duez, Danielle. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25(1): 11–28.
Duez, Danielle. 1985. Perception of silent pauses in spontaneous speech. *Language and Speech* 28(4): 377–89.
Duez, Danielle. 1993. Acoustic correlates of subjective pauses. *Journal of Psycholinguistic Research* 22(1): 21–39.
Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Malden, MA/Oxford: Blackwell.
Eckert, Penelope. 2005. Variation, convention, and social meaning. Paper presented at the Linguistic Society of America 2005 Annual Meeting: Oakland, CA.
Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4): 453–76.
Eckert, Penelope. Under review. Three waves of variation study: the emergence of meaning in the study of variation.
Eckert, Penelope and John R. Rickford (eds). 2001. *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.
Edlund, Jes, Mattias Heldner, and Julia Hirschberg. 2009. Pause and gap length in face-to-face interaction. *Proceedings of Interspeech 2009*, Brighton, UK.
Edwards, Jane. 2001. The transcription of discourse. In Deborah Tannen, Deborah Schiffrin, and Heidi Hamilton (eds), *Handbook of Discourse Analysis*. Malden, MA/Oxford: Blackwell, 321–48.
Evans, Bronwen and Paul Iverson. 2004. Vowel normalization for accent: an investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America* 115: 352–61.
Evans, Bronwen and Paul Iverson. 2007. Plasticity in vowel perception and production: a study of accent change in young adults. *Journal of the Acoustical Society of America* 121: 3814–26.
Fasold, Ralph W. 1972. *Tense Marking in Black English: a Linguistic and Social Analysis*. Washington, DC: Center for Applied Linguistics.
Feagin, Crawford. 1987. A closer look at the Southern drawl: Variation taken to extremes. In Keith Denning, Sharon Inkelas, Faye McNair-Knox, and John R. Rickford (eds), *Variation in Language: NWAV-XV at Stanford*. Palo Alto, CA: Stanford Department of Linguistics, 137–50.

Feldstein, Stanley and Cynthia Crown. 1978. Conversational time patterns as a function of introversion and extraversion. Paper presented at the Eastern Psychological Association: Washington, DC.
Feldstein, Stanley and Barbara Sloan. 1984. Actual and stereotyped speech tempos of extraverts and introverts. *Journal of Personality* 52(2): 188–204.
Feldstein, Stanley, Luciano Alberti, and Mohammed BenDebba. 1979. Self-attributed personality characteristics and the pacing of conversational interaction. In Aron W. Siegman and Stanley Feldstein (eds), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 73–87.
Feldstein, Stanley, Faith-Anne Dohm, and Cynthia Crown. 1993. Gender as a mediator in the perception of speech rate. *Bulletin of the Psychonomic Society* 31(6): 521–4.
Fischer, John L. 1958. Social influences on the choice of a linguistic variant. *Word* 14: 47–56.
Fónagy, Ivan and Klara Magdics. 1960. Speed of utterance in phrases of different lengths. *Language and Speech* 3: 179–92.
Foulkes, Paul and Gerard Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34: 409–38.
Freiman, Howard. 1979. Speech rate as a function of dialect geography. In Robert St. Clair (ed.), *Perspectives on Applied Sociolinguistics: From the Language Medium of Education to the Semiotic Language of the Media*. Lawrence, KS: Coronado Press, 128–36.
Fridland, Valerie and Tyler Kendall. 2012. Exploring the relationship between production and perception in the mid front vowels of US English. *Lingua* 122(7): 779–93.
Froyum Roise, Carissa. 2004. "Doing what I do": African American teenagers, gender, and sexuality in an inner city. MA thesis. Raleigh, NC: North Carolina State University.
Garofolo, John S., Lori Lamel, William Fisher, Jonathan Fiscus, David Pallett, Nancy Dahlgren, and Victor Zue. 1993. TIMIT Acoustic-Phonetic Continuous Speech Corpus. Philadelphia: Linguistic Data Consortium. Corpus.
Geeraerts, Dirk, Gitte Kristiansen, and Yves Peirsman (eds). 2010. *Advances in Cognitive Sociolinguistics*. Berlin/New York: Mouton de Gruyter.
Giles, Howard. 1973. Accent mobility: a model and some data. *Anthropological Linguistics* 15(2): 87–105.
Giles, Howard and Peter Powesland. 1975. *Speech Style and Social Evaluation*. London/New York: Academic Press.
Giles, Howard and Philip M. Smith. 1979. Accommodation theory: Optimal levels of convergence. In Howard Giles and Robert St. Clair (eds), *Language and Social Psychology*. Malden, MA/Oxford: Blackwell, 45–65.
Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. In Howard Giles, Justine Coupland, and Nikolaus Coupland (eds), *Contexts of Accommodation: Developments in Applied Sociolinguistics*, Cambridge: Cambridge University Press, 1–68.
Giles, Howard, Nikolas Coupland, Karen Henwood, Jim Harriman, and Justine Coupland. 1990. The social meaning of RP: an intergenerational perspective. In Susan Ramgaran (ed.), *Studies in the Pronunciation of English: a Commemorative Volume in Honor of A. C. Gimson*. New York: Routledge, 191–211.

Goldman-Eisler, Frieda. 1954. On the variability of the speed of talking and on its relation to the length of utterance in conversation. *British Journal of Psychology* 45(2): 94–107.

Goldman-Eisler, Frieda. 1958. The predictability of words in context and the length of pauses in speech. *Language and Speech* 1(3): 226–31.

Goldman-Eisler, Frieda. 1961. The significance of changes in the rate of articulation. *Language and Speech* 4(4): 171–4.

Goldman-Eisler, Frieda. 1968. *Psycholinguistics: Experiments in Spontaneous Speech*. London/New York: Academic Press.

Gordon, Elizabeth, Margaret Maclagan, and Jennifer Hay. 2007. The ONZE Corpus. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 2: *Diachronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, 82–104.

Gregersen, Frans. 2009. The data and design of the LANCHART study. *Acta Linguistica Hafniensia* 41: 3–29.

Gregersen, Frans and Michael Barner-Rasmussen. 2011. The logic of comparability: On genres and phonetic variation in a project on language change in real time. *Corpus Linguistics and Linguistic Theory* 7(1): 7–36.

Gregersen, Frans, Søren Beck Nielsen, and Jacob Thøgersen. 2009. Stepping into the same river twice: On the discourse context analysis in the LANCHART project. *Acta Linguistica Hafniensia* 41: 30–63.

Gries, Stefan Th. 2009. *Statistics for Linguistics with R: a Practical Introduction*. Berlin: Walter de Gruyter.

Grosjean, François. 1980a. Linguistic structures and performance structures: Studies in pause distribution. In Hans Dechert and Manfred Raupach (eds), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton, 91–106.

Grosjean, François. 1980b. Comparative studies of temporal variables in spoken and sign languages: a short review. In Hans Dechert and Manfred Raupach (eds), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton, 307–12.

Grosjean, François and Alain Deschamps. 1975. Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 31(3–4): 144–84.

Guy, Gregory. 1980. Variation in the group and the individual: the case of final stop deletion. In William Labov (ed.), *Locating Language in Time and Space*. New York: Academic, 1–36.

Guy, Gregory and Cecelia Cutler. 2011. Speech style and authenticity: Quantitative evidence for the performance of identity. *Language Variation and Change* 23: 139–62.

Harrell, Frank E. Jr. 2001. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.

Harrell, Frank E. Jr. 2009. Design: Design Package for R.

Harrell, Frank E. Jr. 2011. rms: Regression Modeling Strategies for R.

Hay, Jennifer, Katie Drager, and Paul Warren. 2009. Careful who you talk to: an effect of experimenter identity on the production of the NEAR/SQUARE merger in New Zealand English. *Australian Journal of Linguistics* 29(2): 269–85.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34: 458–84.

Hay, Jennifer, Paul Warren, and Katie Drager. 2010. Short-term exposure to one dialect affects processing of another. *Language and Speech* 53(4): 447–71.

Hazen, Kirk. 2000a. *Identity and Ethnicity in the Rural South: a Sociolinguistic View through Past and Present Be*. Publication of the American Dialect Society 83. Durham, NC: Duke University Press.

Hazen, Kirk. 2000b. The role of researcher identity in conducting sociolinguistic research: a reflective case study. *Southern Journal of Linguistics* 24(1): 103–15.

Hazen, Kirk. 2008. (ING): a vernacular baseline for English in Appalachia. *American Speech* 83(2): 116–40.

Hazen, Kirk. 2011. Flying high above the social radar: Coronal stop deletion in modern Appalachia. *Language Variation and Change* 23: 105–37.

Henderson, Alan. 1974. Time patterns in spontaneous speech: Cognitive stride or random walk? – A reply to Jaffe et al. (1972). *Language and Speech* 17: 119–25.

Henderson, Alan, Frieda Goldman-Eisler, and Andrew Skarbek. 1966. Sequential temporal patterns in spontaneous speech. *Language and Speech* 9: 207–16.

Hewlett, Nigel and Monica Rendall. 1998. Rural versus urban accent as an influence on the rate of speech. *Journal of the International Phonetic Association* 28(1–2): 63–71.

Hymes, Dell. 1974. *Foundations in Sociolinguistics: an Ethnographic Approach*. Philadelphia: University of Pennsylvania.

Jacewicz, Ewa, Robert A. Fox, Caitlin O'Neill, and Joseph Salmons. 2009. Articulation rate across dialect, age, and gender. *Language Variation and Change* 21(2): 233–56.

Jacewicz, Ewa, Robert A. Fox, and Lai Wei. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America* 128(2): 839–50.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59: 434–46.

Jaffe, Joseph and Stephen Breskin. 1970. Temporal patterns of speech and sample size. *Journal of Speech and Hearing Research* 13(3): 667–8.

Jaffe, Joseph and Stanley Feldstein. 1970. *Rhythms of Dialogue*. New York: Academic Press.

Jaffe, Joseph, Stephen Breskin, and Louis J. Gerstman. 1972. Random generation of apparent speech rhythms. *Language and Speech* 15: 68–71.

Johnson, Daniel Ezra. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3(1): 359–83.

Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. Malden, MA/Oxford: Blackwell.

Johnstone, Barbara. 2007. *Discourse Analysis*, 2nd edn. Malden, MA/Oxford: Blackwell.

Kapatsinski, Vsevelod. 2010. What is it I am writing? Lexical frequency effects in spelling Russian prefixes: Uncertainty and competition in an apparently regular system. *Corpus Linguistics and Linguistic Theory* 6(2): 157–215.

Kendall, Tyler. 2005. Advancing the utility of the transcript: a computer-enhanced methodology. Paper presented at the Twelfth International Conference on Methods in Dialectology: Moncton, New Brunswick, Canada.

Kendall, Tyler. 2006. A sociolinguistic analysis of pause in North Carolina. Paper presented at the South Eastern Conference on Linguistics (SECOL) 73: Auburn, Alabama.

Kendall, Tyler. 2006–2007. Advancing the utility of the transcript: a computer-enhanced methodology. *Linguistica Atlantica* 27–28: 51–5.

Kendall, Tyler. 2007a. The Sociolinguistic Archive and Analysis Project: Empowering the sociolinguistic archive. *Penn Working Papers in Linguistics* 13(2): 15–26.

Kendall, Tyler. 2007b. On the status of pause in sociolinguistics. Paper presented at the Linguistic Society of America 2007 Annual Meeting: Anaheim, CA.

Kendall, Tyler. 2008a. On the history and future of sociolinguistic data. *Language and Linguistics Compass* 2(2): 332–51.

Kendall, Tyler. 2008b. Identity, performance, and "consciousness": the use of locally salient linguistic forms in a formerly isolated community. Paper presented at New Ways of Analyzing Variation (NWAV) 37: Houston, TX.

Kendall, Tyler. 2009. Speech rate, pause, and linguistic variation: an examination through the Sociolinguistic Archive and Analysis Project. Doctoral dissertation. Durham, NC: Duke University.

Kendall, Tyler. 2010a. Developing web interfaces to spoken language data collections. *Proceedings of the Chicago Colloquium on Digital Humanities and Computer Science* 1(2). https://letterpress.uchicago.edu/index.php/jdhcs/article/view/53

Kendall, Tyler. 2010b. Accommodating (ING): Individual variation in mixed-ethnicity interviews. In Barry Heselwood and Clive Upton (eds), *Proceedings of Methods XIII: Papers from the Thirteenth International Conference on Methods in Dialectology, 2008*. Frankfurt am Main: Peter Lang, 351–61.

Kendall, Tyler. 2011. Corpora from a sociolinguistic perspective (Corpora sob uma perspectiva sociolinguística). In Stefan Th. Gries (ed.), Corpus studies: Future directions, special issue of *Revista Brasileira de Linguística Aplicada* 11(2): 361–89.

Kendall, Tyler. forthcoming a. Data preservation and access. In Christine Mallinson, Becky Childs, and Gerard Van Herk (eds), *Data Collection in Sociolinguistics: Methods and Applications*. New York/London: Routledge.

Kendall, Tyler. forthcoming b. Data in the study of language variation and change. In J. K. Chambers and Natalie Schilling (eds), *The Handbook of Language Variation and Change*, 2nd edn. Malden, MA/Oxford: Wiley-Blackwell.

Kendall, Tyler and Ann R. Bradlow. 2011. Mobilizing smaller datasets for large-scale phonetic analysis: Web-databases and semi-automatic analyses. Paper presented at New Tools and Methods for Very-Large-Scale Phonetics Research: Philadelphia, PA.

Kendall, Tyler and Valerie Fridland. 2012. Variation in perception and production of mid front vowels in the US Southern Vowel Shift. *Journal of Phonetics* 40: 289–306.

Kendall, Tyler and Christine Mallinson. 2008. "Narrative style" vs. "interview style" in the talk of African American girls in Washington, DC. Paper presented at Georgetown University Round Table (GURT) 2008: Washington, DC.

Kendall, Tyler and Erik R. Thomas. 2010. Dissecting rate of speech: the effect of phrase final lengthening on articulation rate. Poster presented at the Second Pan American/Iberian Meeting on Acoustics/160th Meeting of the Acoustical Society of America (ASA160): Cancun, Mexico.

Kendall, Tyler and Gerard Van Herk. 2011. Corpus linguistics and sociolinguistic inquiry: Introduction to special issue. *Corpus Linguistics and Linguistic Inquiry* 7(1): 1–6.

Kircher, Tilo, Michael Brammer, Willem Levelt, Mathias Bartels, and Philip McGuire. 2004. Pausing for thought: Engagement of left temporal cortex during pauses in speech. *NeuroImage* 22: 84–90.

Kowal, Sabine and Daniel O'Connell. 1980. Pausological research at Saint Louis University. In Hans Dechert and Manfred Raupach (eds), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*. The Hague: Mouton, 61–6.

Kowal, Sabine and Daniel O'Connell. 1985. Cognitive rhythms reluctantly revisited. Letter to the editor. *Language and Speech* 28: 93–5.

Kowal, Sabine, Daniel O'Connell, and Edward Sabin. 1975. Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research* 4(3): 195–207.

Kristiansen, Gitte and René Dirven (eds). 2008. *Cognitive Sociolinguistics: Langauge Variation, Cultural Models, Social Systems*. Berlin/New York: Mouton de Gruyter.

Labov, William. 1963. The social motivation of a sound change. *Word* 19: 273–309.

Labov, William. 1966[2006]. *The Social Stratification of English in New York City*, 2nd edn. Cambridge: Cambridge University Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.

Labov, William. 1994. *Principles of Linguistic Change*, Vol. 1: *Internal Factors*. Malden, MA/Oxford: Blackwell.

Labov, William. 2001. *Principles of Linguistic Change*, Vol. 2: *Social Factors*. Malden, MA/Oxford: Blackwell.

Labov, William. 2010. *Principles of Linguistic Change*, Vol. 3: *Cognitive and Cultural Factors*. Malden, MA/Oxford: Wiley-Blackwell.

Labov, William and Joshua Waletzky. 1967. Narrative analysis. In June Helm (ed.), *Essays on the Verbal and Visual Arts*. Seattle: University of Washington, 12–44.

Labov, William, Sharon Ash and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: De Gruyter.

Ladefoged, Peter. 2006. *A Course in Phonetics*, 5th edn. Boston: Thomson Wadsworth.

Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lay, Clarry H. and Allan Paivio. 1969. The effects of task difficulty and anxiety on hesitations in speech. *Canadian Journal of Behavioral Science* 1: 25–37.

Lehtonen, Jaako. 1979. Speech rate and pauses in the English of Finns, Swedish-speaking Finns, and Swedes. In Rolf Palmberg (ed.), *Perception and Production of English: Papers on Interlanguage*. Åbo: Åbo Akademi, 35–51.

Lehtonen, Jaako and Kari Sajavaara. 1985. The silent Finn. In Deborah Tannen and Muriel Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, 193–201.

Levelt, Willem. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levin, H. and I. Silverman. 1965. Hesitation phenomena in children's speech. *Language and Speech* 8: 67–85.

Liddicoat, Anthony. 2007. *An Introduction to Conversation Analysis*. London/New York: Continuum.
Luke, Douglas A. 2004. *Multilevel Modeling*. Quantitative Applications in the Social Sciences 143. Thousand Oaks, CA: Sage Publications.
Macaulay, Ronald. 2002. Discourse variation. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Malden, MA/Oxford: Blackwell, 283–305.
McEnery, Tony and Andrew Wilson. 2001. *Corpus Linguistics*, 2nd edn. Edinburgh: Edinburgh University Press.
McEnery, Tony, Richard Xiao, and Yukio Tono. 2006. *Corpus-based Language Studies: an Advanced Resource Book*. New York/London: Routledge.
Maclay, Howard and Charles E. Osgood. 1959. Hesitation phenomena in spontaneous speech. *Word* 15: 19–44.
MacWhinney, Brian. 2007. The TalkBank Project. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, 163–80.
Mallinson, Christine and Becky Childs. 2007. Communities of practice in sociolinguistic description: Analyzing language and identity practices among black women in Appalachia. *Gender and Language* 1(2): 173–206.
Mallinson, Christine and Tyler Kendall. 2009. "The way I can speak for myself": the social and linguistic context of counseling interviews with African American adolescent girls in Washington, DC. In Sonja Lanehart (ed.), *African American Women's Language*. Newcastle upon Tyne: Cambridge Scholars Press, 110–26.
Mallinson, Christine and Walt Wolfram. 2002. Dialect accommodation in a bi-ethnic mountain enclave community: More evidence on the development of African American English. *Language in Society* 31: 743–75.
Martin, James G. and Winifred Strange. 1968. The perception of hesitation in spontaneous speech. *Perception and Psychophysics* 3(6): 427–38.
Mendoza-Denton, Norma. 1995. Pregnant pauses: Silence and authority in the Anita Hill–Clarence Thomas hearings. In Kira Hall and Mary Bucholtz (eds), *Gender Articulated: Language and the Socially Constructed Self*. New York/London: Routledge, 51–66.
Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Malden, MA/Oxford: Blackwell.
Miller, Joanne L., François Grosjean, and Concetta Lomato. 1984. Articulation rate and its variability in spontaneous speech: a reanalysis and some implications. *Phonetica* 41(4): 215–25.
Miller, Norman, Geoffrey Maruyama, Rex J. Beaber, and Keith Valone. 1976. Speed of speech and persuasion. *Journal of Personality and Social Psychology* 34: 615–25.
Milroy, Lesley. 1987. *Observing and Analyzing Natural Language*. Malden, MA/Oxford: Blackwell.
Milroy, Lesley and Matthew Gordon. 2003. *Sociolinguistics: Methods and Interpretation*. Malden, MA/Oxford: Blackwell.
Moreton, Elliott, and Erik Thomas. 2007. Origins of Canadian raising in voiceless-coda effects: a case study in phonologization. In Jennifer Cole and José I. Hualde (eds), *Papers in Laboratory Phonology 9*. Berlin/New York: Mouton de Gruyter, 37–63.

Mukherjee, Joybrato. 2000. Speech is silver, but silence is golden: Some remarks on the function(s) of pauses. *Anglia* 118: 571–84.
Niedzielski, Nancy and Dennis Preston. 2003. *Folk Linguistics*. Berlin: Walter de Gruyter.
O'Connell, Daniel and Sabine Kowal. 1972. Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research* 1(2): 155–64.
Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs and Bambi Schieffelin (eds), *Developmental Pragmatics*. New York: Academic Press, 43–72.
Pardo, Jennifer. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119(4): 2382–93.
Philips, Susan U. 1976. Some sources of cultural variability in the regulation of talk. *Language in Society* 5(1): 81–95.
Pinheiro, José C. and Douglas M. Bates. 2000. *Mixed-Effect Models in S and S-PLUS*. New York: Springer.
Podesva, Robert. 2007. Phonation type as a stylistic variable: the use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11: 478–504.
Poplack, Shana. 1989. The care and handling of a mega-corpus: the Ottawa–Hull French Project. In Ralph Fasold and Deborah Schiffrin (eds), *Language Change and Variation*. Amsterdam: John Benjamins, 411–51.
Poplack, Shana. 2007. Foreward. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, ix–xiii.
Power, M. J. 1983. Are there cognitive rhythms in speech? *Language and Speech* 26: 253–61.
Power, M. J. 1984. Are there cognitive rhythms in speech? – A reply to Beattie (1984). *Language and Speech* 27: 197–200.
Preston, Dennis. 1982. 'Ritin fowklower daun 'rong: Folklorists' failures in phonology. *Journal of American Folklore* 95: 304–26.
Preston, Dennis. 1989. *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics*. Dordrecht: Foris Publications.
Preston, Dennis. 1999. A language attitude approach to the perception of regional variation. In Dennis Preston (ed.), *The Handbook of Perceptual Dialectology*, Vol. 1. Amsterdam: John Benjamins, 359–73.
Preston, Dennis. 2000. Mowr and mowr bayud spellin': Confessions of a sociolinguist. *Journal of Sociolinguistics* 4(4): 614–21.
Preston, Joan and Robert Gardner. 1967. Dimensions of oral and written language fluency. *Journal of Verbal Learning and Verbal Behavior* 6: 936–45.
Quené, Hugo. 2007. On the just noticeable difference for tempo in spontaneous speech. *Journal of Phonetics* 35: 353–62.
Quené, Hugo. 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *Journal of the Acoustical Society of America* 123(2): 1104–13.
Quené, Hugo and Huug van den Bergh. 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59: 413–25.
R Development Core Team. 2011. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Ramsay, R. W. 1968. Speech patterns and personality. *Language and Speech* 11(1): 54–63.

Ray, George and Christopher Zahn. 1990. Regional speech rates in the United States: a preliminary analysis. *Communication Research Reports* 7: 34–7.

Ray, George and Christopher Zahn. 1999. Language attitudes and speech behavior: New Zealand English and Standard American English. *Journal of Language and Social Psychology* 18(3): 310–19.

Ray, George, Eileen Ray, and Christopher Zahn. 1991. Speech behavior and social evaluation: an examination of medical messages. *Communication Quarterly* 39(2): 119–29.

Redford, Melissa. 1999. An articulatory basis for the syllable. Doctoral dissertation. Austin, TX: University of Texas at Austin.

Redford, Melissa. In press. A comparative analysis of pausing in child and adult storytelling. *Applied Psycholinguistics.*

Reynolds, Allan and Allan Paivio. 1968. Cognitive and emotional determinants of speech. *Canadian Journal of Psychology* 22: 164–75.

Rickford, John R. and Faye McNair-Knox. 1994. Addressee- and topic-influenced style shift: a quantitative sociolinguistic study. In Douglas Biber and Edward Finegan (eds), *Perspectives on Register: Situating Register Variation within Sociolinguistics*. New York/Oxford: Oxford University Press, 235–76.

Roach, Peter. 1998. Myth 18: Some languages are spoken more quickly than others. In Laurie Bauer and Peter Trudgill (eds), *Language Myths*. London/New York: Penguin, 150–8.

Robb, Michael, Margaret Maclagan, and Yang Chen. 2004. Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics and Phonetics* 18: 1–15.

Rochester, S. R. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research* 2(1): 51–81.

Sabin, Edward J., Edward J. Clemmer, Daniel C. O'Connell, and Sabine Kowal. 1979. A pausological approach to speech development. In Aron W. Siegman and Stanley Feldstein (eds), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 35–55.

Sajavaara, Kari and Jaakko Lehtonen. 1997. The silent Finn revisited. In Adam Jaworski (ed.), *Silence: Interdisciplinary Perspectives*. Berlin: Mouton de Gruyter, 263–83.

Sankoff, David and Suzanne Laberge. 1978. The linguistic market and the statistical explanation of variability. In David Sankoff (ed.), *Linguistic Variation: Models and Methods*. New York: Academic Press, 239–50.

Sankoff, David, Sali Tagliamonte, and Eric Smith. 2005. Goldvarb X: a variable rule application for Macintosh and Windows. Toronto: Dept of Linguistics, University of Toronto.

Sankoff, Gillian and Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83: 560–88.

Saville-Troike, Muriel. 1985. The place of silence in an integrated theory of communication. In Deborah Tannen and Muriel Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, 3–18.

Schilling-Estes, Natalie. 1998. Investigating "self-conscious" speech: the performance register in Ocracoke English. *Language in Society* 27: 53–83.

Schilling-Estes, Natalie. 2002. Investigating stylistic variation. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Malden, MA/Oxford: Blackwell, 375–401.

Schilling-Estes, Natalie. 2004. Constructing ethnicity in interaction. *Journal of Sociolinguistics* 8(2): 163–95.

Schnoebelen, Tyler. 2009. The social meaning of tempo. Unpublished manuscript. Palo Alto, CA: Stanford University.

Schnoebelen, Tyler. 2010. Variation in speech tempo: Capt. Kirk, Mr. Spock, and all of us in between. Paper presented at New Ways of Analyzing Variation (NWAV) 39. San Antonio, TX.

Searle, John. 1969. *Speech Acts: an Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Shannon, Claude E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423, 623–56.

Shuy, Roger, Walt Wolfram, and William Riley. 1968. *Field Techniques in an Urban Language Study*. Washington, DC: Center for Applied Linguistics.

Siegman, Aron W. 1978. The meaning of silent pause in the interview. *The Journal of Nervous and Mental Disease* 166: 642–54.

Siegman, Aron W. 1979a. The voice of attraction: Vocal correlates of interpersonal attraction in the interview. In Aron W. Siegman and Stanley Feldstein (eds), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 89–113.

Siegman, Aron W. 1979b. Cognition and hesitation in speech. In Aron W. Siegman and Stanley Feldstein (eds), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 151–78.

Siegman, Aron W. and Mark Reynolds. 1982. Interviewer–interviewee nonverbal communications: an interactional approach. In Martha A. Davis (ed.), *Interaction Rhythms: Periodicity in Communication Behavior*. New York: Human Sciences Press, 249–78.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23(3–4): 193–229.

Sledd, James H. 1966. Breaking, umlaut, and the Southern drawl. *Language* 42(1): 18–41.

Smith, Bruce L., Bruce L. Brown, William J. Strong, and Alvin C. Rencher. 1975. Effects of speech rate on personality perception. *Language and Speech* 18(2): 145–52.

Staum Casasanto, Laura, Kyle Jasmin, and Daniel Casasanto. 2010. Virtually accommodating: Speech rate accommodation to a virtual interlocutor. In Stellan Ohlsson and Richard Catrambone (eds), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 127–32.

Strand, Elizabeth and Keith Johnson. 1996. Gradient and visual speaker normalization in the perception of fricatives. In Dafydd Gibbon (ed.), *Natural Language Processing and Speech Technology*. Berlin: Mouton, 14–26.

Strangert, Eva. 1993. Speaking style and pausing. In Eva Strangert, Mattias Heldner and Peter Czigler (eds), *Reports from the Department of Phonetics, University of Umeå (PHONUM) 2*. Umeå, Sweden: University of Umeå, 121–37.

Street, Richard L., Jr and Robert M. Brady. 1982. Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs* 49(4): 290–308.

Street, Richard L., Jr, Robert M. Brady, and William B. Putman. 1983. The influence of speech rate stereotypes and rate similarity on listeners' evaluations of speakers. *Journal of Language and Social Psychology* 2(1): 37–56.
Street, Richard L., Jr., Robert M. Brady, and Raymond Lee. 1984. Evaluative responses to communicators: the effects of speech rate, sex, and interaction context. *Western Journal of Speech Communication* 48: 14–27.
Sumner, Meghan and Arthur Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60: 487–501.
Tagliamonte, Sali. 2004. Somethi[ŋ]'s goi[n] on!: Variable (ing) at ground zero. In Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fidell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundergren, and Mats Thelander (eds), *Language Variation in Europe: Papers from the Second International Conference on Language Variation in Europe, ICLaVE 2*. Uppsala, Sweden: Uppsala University, 390–403.
Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
Tagliamonte, Sali. 2007. Representing real language: Consistency, trade-offs and thinking ahead! In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora*. Vol. 1: *Synchronic Databases*. New York/Basingstoke, Hampshire: Palgrave Macmillan, 205–40.
Tagliamonte, Sali. 2008. Conversations from the speech community: Exploring language variation in synchronic dialect corpora. In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta, and Minna Korhonen (eds), *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*. Amsterdam: John Benjamins, 107–28.
Tannen, Deborah. 1984[2005]. *Conversational Style: Analyzing Talk among Friends*, rev. edn. New York/Oxford: Oxford University Press.
Tannen, Deborah. 1985. Silence: Anything but. In Deborah Tannen and Muriel Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, 93–111.
Tannen, Deborah. 2000. "Don't just sit there – interrupt!": Pacing and pausing in conversational style. *American Speech* 75(4): 393–5.
Tannen, Deborah and Muriel Saville-Troike (eds). 1985. *Perspectives on Silence*. Norwood, NJ: Ablex.
Thakerar, Jitendra N. and Howard Giles. 1981. They are – so they speak: Noncontent speech stereotypes. *Language and Communication* 1: 251–6.
Thakerar, Jitendra N., Howard Giles, and Jenny Cheshire. 1982. Psychological and linguistic parameters of speech accommodation theory. In Colin Fraser and Klaus R. Scherer (eds), *Advances in the Social Psychology of Language*. Cambridge: Cambridge University Press, 205–55.
Thomas, Erik R. 2002a. Instrumental phonetics. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Oxford/Malden, MA: Blackwell, 168–200.
Thomas, Erik R. 2002b. Sociophonetic applications of speech perception experiments. *American Speech* 77(2): 115–47.
Thomas, Erik R. 2010. A longitudinal analysis of the durability of the Northern/Midland boundary in Ohio. *American Speech* 85(4): 375–430.
Thomas, Erik R. 2011a. *Sociophonetics: an Introduction*. New York/Basingstoke, Hampshire: Palgrave Macmillan.

Thomas, Erik R. 2011b. Sociolinguistic variables and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*. http://dx.doi.org/10.1002/wcs.152
Thomas, Erik R. and Phillip M. Carter. 2006. Prosodic rhythm and African American English. *English World-Wide* 27(3): 331–55.
Thomas, Erik R. and Holly Ericson. 2007. Intonational distinctiveness of Mexican American English. *Penn Working Papers in Linguistics* 13(2): 193–205.
Torbert, Benjamin. 2001. Language history traced through consonant cluster reduction: the case of Native American Lumbee English. *American Speech* 75: 361–87.
Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1: 179–95.
Tsao, Ying Chiao and Gary Weisner. 1997. Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research* 40(4): 858–66.
Turk, Alice and Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35: 445–72.
Vaane, Eveline. 1982. Subjective estimation of speech rate. *Phonetica* 39: 136–49.
Van Borsel, John and Dorothy De Maesschalck. 2008. Speech rate in males, females, and male-to-female transsexuals. *Clinical Linguistics and Phonetics* 22(9): 679–85.
Van Engen, Kristin J., Melissa Baese-Berk, Rachel E. Baker, Arim Choi, Midam Kim, and Ann R. Bradlow. 2010. The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech* 53(4): 510–40.
Weinreich, Uriel, William Labov, and Marvin Herzog. (1968). Empirical foundations for a theory of language change. In Winfred P. Lehmann and Yakov Malkiel (eds), *Directions for Historical Linguistics*. Austin, TX: University of Texas Press, 95–195.
Wells, John C. 1982. *Accents of English: an Introduction*. Cambridge: Cambridge University Press.
Wetzell, W. Brett. 2000. Rhythm, dialects, and the Southern drawl. MA thesis. Raleigh, NC: North Carolina State University.
White, Laurence and Sven Mattys. 2007. Rhythmic typology and variation in first and second languages. In Pilar Prieto, Joan Mascaró, and Maria-Josep Solé (eds), *Segmental and Prosodic Issues in Romance Phonology*. Amsterdam: John Benjamins, 237–57.
Wolford, Tonya and Phillip M. Carter. 2007. Dying a slow death: Tracing the complexities of language use, loss, and persistence in the Spanish of South Texas. Paper presented at New Ways of Analyzing Variation (NWAV) 36: Philadelphia, PA.
Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.
Wolfram, Walt. 1993. Identifying and interpreting variables. In Dennis Preston (ed.), *American Dialect Research*. Amsterdam: John Benjamins, 193–221.
Wolfram, Walt. 1999. *Map of North Carolina Dialects*. Raleigh, NC: North Carolina State University.
Wolfram, Walt. 2006. Variation in language: Overview. In Keith Brown (ed.), *Encyclopedia of Languages and Linguistics*, 2nd edn. Oxford: Elsevier, 333–40.

Wolfram, Walt and Natalie Schilling-Estes. 1995. Moribund dialects and the endangerment canon: the case of the Ocracoke brogue. *Language* 71: 696–721.
Wolfram, Walt and Erik R. Thomas. 2002. *The Development of African American English*. Oxford, UK/ Malden, MA: Blackwell.
Wolfram, Walt, Becky Childs, and Benjamin Torbert. 2000. Tracing English dialect history through consonant cluster reduction: Comparative evidence from isolated dialects. *Southern Journal of Linguistics* 24: 17–40.
Wolfram, Walt, Clare Dannenberg, Stanley Knick, and Linda Oxendine. 2002. *Fine in the World: Lumbee Language in Time and Place*. Raleigh, NC: North Carolina State University.
Yuan, Jiahong and Mark Liberman. 2008. Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics 2008*.
Yuan, Jiahong, Mark Liberman, and Christopher Cieri. 2006. Towards an integrated understanding of speaking rate in conversation. Paper presented at the Ninth International Conference on Spoken Language Processing (Interspeech 2006): Pittsburgh, Pennsylvania.
Zellner, Brigitte. 1994. Pauses and the temporal structure of speech. In Eric Keller (ed.), *Fundamentals of Speech Synthesis and Speech Recognition*. Chichester: Wiley, 41–62.
Zhang, Qing. 2008. Rhotacization and the "Beijing Smooth Operator": the social meaning of a linguistic variable. *Journal of Sociolinguistics* 12: 201–22.

Index