12 Transcription

Naomi Nagy and Devyani Sharma

> To write Faetar, you have to use the Italian spelling system, because it's the only system the speakers know.
>
> When you write Faetar, you should use French orthography because that will indicate the Gallic roots of the language.
>
> Of course, the only option is to use the International Phonetic Alphabet to write Faetar, so that linguists around the world are able to understand the details of our unique language.

1 Introduction

The reconstructed vignette above, based on actual conversations with speakers of Faetar, an endangered language spoken in two small villages in southern Italy (Nagy 2000, 2011a), illustrates some of the many uses that transcription has. Researchers (and the transcribers they hire) may not even be aware of all the potential downstream uses of their transcriptions. The most common understanding of the purpose of transcription in linguistics is contained in the third statement. However, a linguist's decision to transcribe in a standardized orthography or in the International Phonetic Alphabet (IPA) can influence later uses of the text. Deviations from the traditions of one's field can even be perceived as ideologically charged. As Kendall (2008: 337) puts it,

> the act of transcription [. . .] is often undertaken as a purely methodological activity, as if it were theory neutral. Each decision that is made while transcribing influences and constrains the resulting possible readings and analyses (Ochs 1979; Mishler 1991; Bucholtz 2000; Edwards 2001). Decisions as seemingly straightforward as how to lay out the text, to those more nuanced – like how much non-verbal information to include and how to encode minutiae such as pause length and utterance overlap – have far-reaching effects on the utility of a transcript and the directions in which the transcript may lead analysts.

Acknowledgments: We are very grateful to the following people who shared their transcribing experiences, expertise, and protocols: Julie Auger, Jenny Cheshire, Ashwini Deo, Nathalie Dion, John Du Bois, Sue Fox, Alexandra Georgakopoulou, Kirk Hazen, Dagmar Jung, Tyler Kendall, Shana Poplack, Ben Rampton, Rena Torres Cacoullos, Albert Valdman, Anne-José Villeneuve, and Walt Wolfram. All personal communications cited in the text are from July 2011.

Transcription can serve a wide range of functions, as a single transcript may eventually be used for multiple analyses. Within linguistic research, a transcript may be used, for instance, for quantitative analysis of morphosyntactic or discourse variables, as a guide for auditory phonetic analysis, for qualitative analysis of conversation, discourse, or interaction, and for theoretical linguistic analysis. In addition to serving linguistic research, the transcriptions may become a legacy, providing documentation of a particular point in a language variety's development, as well as recording information about the culture of the society who used the variety. Often, only transcripts (not accompanied by the recordings they represent) are shared with the public and other researchers, making their accuracy critical, as they must represent everything deemed important from the original recording. Transcripts might even be used by the community long after research is completed: Transcripts from Walt Wolfram's research (pers. comm.) have been used by members of the community to compile oral history CDs (e.g., *Ocracoke Speaks 2001*), and communities have even asked the researchers on that project to transcribe tapes for them. These many different needs and uses set different requirements for transcription practices and protocols.

In this chapter, we discuss various dimensions of two broad questions: what to transcribe and how to transcribe it, what Bucholtz (2000) terms "interpretive" and "representational" decisions respectively. The chapter breaks these two dimensions down to cover a range of issues: aspects of form and content when transcribing, transcribing across languages, the advantages of different types of software in transcription, transcriber effects, transcription protocols, and practicalities of planning transcription.

2 How much to transcribe

Although we may think of transcription as a more or less mechanical "translation" from an oral medium to a text medium, there are in fact many decisions that must be made regarding what parts of a recording to include and what level of detail to indicate for those segments.

The very first decision is whether the voice recording needs to be transcribed at all. Some researchers find it more efficient to proceed directly to extracting the relevant materials or examples from the audio stream, and either transcribe only the relevant passage or code directly without transcribing (Labov's course on sociolinguistic fieldwork methodology in Philadelphia has used this approach). Other linguists prefer to transcribe first so that all material is available in text form. This minimizes the likelihood of missing certain examples and, in the long run, may save time if the recording will be used for a variety of purposes; the Ottawa-Hull Corpus (Poplack 1989), Sankoff and Thibault's Montreal Anglophone project (Sankoff et al. 1997), and the Montreal Francophone Corpora (Sankoff and Sankoff 1973) use this approach. In some such cases, particularly in sociolinguistics, standard practices are employed for the

selection of segments to be transcribed – for instance, omission of the first 10 minutes of a sociolinguistic interview (to avoid speech produced during the less comfortable initial stages of a recording session), or selection of more and less formal speech segments from certain interview topics for stylistic analysis (Labov 2001).

At times, researchers may try to optimize how much can be transcribed by "farming out" the work to professional transcribers. Here, it is worth bearing in mind that linguists are never as close to their object of study as when they are transcribing. The very act of transcription helps the researcher find and understand patterns in the data, seeing elements that may be elusive and fleeting in the original oral form. For this reason, many linguists feel that it is crucial to transcribe as much of their own data as possible.

It is common for researchers simply not to have the funding or time to transcribe a portion of the data collected, particularly with time-consuming bilingual transcription. In some such cases, alternative analyses that permit very limited transcription or auditory processing of data are pursued. In others, only a portion is transcribed. For instance, funding restrictions in Sharma's Dialect Development and Style project meant that recordings from only forty-two of seventy-five individuals could initially be transcribed and analyzed (Sharma and Sankaran 2011); in this case, care had to be taken to select a balanced subsample from each demographic group to avoid skewing in the transcribed portion. In yet other cases, some of the original oral data may not be transcribed simply because more was collected than necessary. For example, the Heritage Language Variation and Change (HLVC) project (Nagy 2009, 2011b) compares speech across forty speakers in each of six languages, and has targeted one hour of transcribed conversational speech as sufficient to represent each speaker. In these cases, decisions must be made about which portions of an interview that exceeds one hour should be included. In the case of the HLVC project, the analysis of sociolinguistic variables focuses on data from 15 minutes into the interview onward, but demographic information about speakers is extracted from any portion of the recording. Therefore, transcription begins at the beginning of the recording, but after the first hour is transcribed, transcribers select only the portions they expect to be useful for demographic description for partial transcription.

Finally, it can happen that portions are not transcribed because they cannot be understood. This is more often the case if the researcher is not a native speaker of the language being studied, as is often the case in endangered language documentation. Sometimes, elements of the context that made understanding possible at the time of the utterance (e.g., gestures or off-microphone interactions) are not recoverable when transcribing. Ambient noise may also make it impossible to determine exactly what was said. In Nagy's experience documenting Faetar, she first transcribed and translated a first draft herself, and then went over any unclear sections with a native speaker. Because speakers, upon seeing a transcript of their own speech, often wish to improve upon what they are recorded as saying, assistants were sought who were not the original speaker.

The question of how much to transcribe extends to elements of content as well. Researchers must make ethical decisions regarding how much personal information to transcribe and how to respect the anonymity often promised to research participants (see Chapter 1). Names of speakers and individuals mentioned are often excluded (or pseudonyms substituted), but further identifying information may also need to be eliminated or altered. Fox and Cheshire (2011) distinguish between *allowable* and *anonymized* references in their Multicultural London English project; examples of both are provided in (1).

(1) Allowable vs anonymized references

Allowable

Havering (a borough where their research was conducted)
I buy my jeans in Mare Street. (general sense of street name)
I used to work in a bar down near Liverpool Street. (general sense of street name)
I'm from E8. (postcode area)

Anonymized

My name's (name of speaker) and I live in Hackney.
I live in (name of street).
I go to (name of school).
if you play football with us yeh over (name of park). (specific places when describing an event)
some white girl from your area. she goes (name of school) she knows (name of girl). (references to schools that could lead to the identity of an individual)
I hated Miss (name of teacher).
Any private information e.g., phone numbers, addresses, specific clubs attended

They also suggest that we carefully consider whether to include references to sexual orientation, date of birth, and "public" individuals such as locally known musicians. It is not always possible to know what information may identify the speaker, depending on the audience, and researchers differ in their views of their obligation to protect the anonymity of speakers who have agreed to be recorded for research purposes. This issue is discussed thoughtfully in Childs, Van Herk, and Thorburn (2011: 176).

Once the relevant segments to be transcribed have been identified, the researcher faces the immediate question of how closely and faithfully to represent the linguistic forms contained in these segments. The sections that follow explore the principles underlying some of these choices.

3 Orthographic choices

Transcription serves as a tool, a "handle" for the original oral recordings, both during primary analysis and for later uses of the data, which may be

years later and not necessarily by the original researcher. Therefore, a well-documented, transparent, and reflexive *orthographic system* is crucial.

The most precise system for transcription is the *International Phonetic Alphabet (IPA)*; see Ladefoged and Disner (2012) for details. This system, usually described as *phonetic transcription* as opposed to other *orthographic transcription*, can potentially render almost all phonetic details of recorded speech faithfully, which may be of crucial importance if the aim of transcription is, for instance, language description or documentation (see Chapter 4). Selective use of IPA may be employed in a transcript if specific dialect variation or contextualization cues are being tracked in a stretch of recorded speech. However, at some point, the time and labor costs of transcribing in IPA must be balanced against the quantity of data to be transcribed and the goals of the research.

In a seminal sociolinguistic research report, Poplack (1989: 430) summarized this orthography issue as follows:

> In planning the transcription of a computer corpus, there is a major trade-off between size of the data base and level of detail of the transcription. For syntactic and lexical work especially, the larger the corpus the better, with the point of diminishing returns nowhere in sight, since a large number of interesting constructions and forms (e.g., most loan-words) are exceedingly rare in natural speech. However, massive corpus size renders fine phonetic transcription unfeasible. Too much detail tends to sharply diminish the utility of automated treatment of the corpus since conventional alphabetical order is lost, and lexically identical forms may be ordered in many different positions.

Given these concerns, the Ottawa-Hull French Project adopted an orthographic approach, rather than phonetic (ibid., p. 431). Pronunciation of particular phonemes was not specified, though omission of entire morphemes was represented by *Ø*, and English borrowings are spelled in English, even when incorporated into French morphology, such as *feeler* and *meaner* (ibid. pp. 432–3).

It is worth noting here that even if phonetic orthography is technically dispensed with – in cases where the transcript is to be used for syntactic analysis, for instance – the transcriber must be alert to phonetic distinctions in order to make orthographic judgments. In Bresnan, Deo, and Sharma (2007), the phonetically "faithful" transcription of verb forms in the *Survey of English Dialects* (Orton et al. 1962–71) were converted to a smaller set of lexical classes that formed the basis of the syntactic analysis of variation in *be*, but fine phonetic distinctions were important in determining the lexical classification of forms. Similarly, in the analysis of syntactic features such as copula omission (e.g., Labov 1969), phonetic reduction of *are* to either *'r* or *Ø* must be extremely carefully coded during transcription, as any phonetic trace of the form is crucial for the outcome of the quantitative syntactic analysis.

The field is now moving toward greater use of *time-aligned transcription* (i.e., textual representations that match stretches of recorded media). Such transcription (currently produced by software discussed later) has multiple advantages: researchers can easily access the original audio(-visual) segment associated with a particular stretch of text; the software usually allows for customized tiers for further interlinear

glossing and tagging of the extract; and data can easily be converted to useful and integrated display formats for presentation (see Thieberger and Berez 2012).

In time-aligned transcription, more "standard" orthography, as opposed to phonetic transcription, is increasingly useful, as phonetic detail may be visualized or easily coded where relevant in later passes through the data. This use of standard orthography also makes computer-assisted analysis easier as different transcribers are less likely to transcribe things differently. As Edwards (2001: 324) observes, "(f)or purposes of computer manipulation (e.g. search, data exchange, or flexible reformatting), the single most important design principle is that *similar instances be encoded in predictably similar ways*" (original emphasis). No matter what specific decisions are made, they should be recorded in a *transcription protocol* that is shared with all researchers using a particular corpus (details are discussed in Section 7).

In the case of morphosyntax, a small set of variants can optionally be agreed upon, noted in the transcription protocol, and used in the transcription (e.g., *ain't*). In the case of phonetic variation and the rendering of connected speech, this is much less common. Researchers generally avoid the use of *eye dialect* (i.e., the use of folk orthographic representations to indicate non-standard pronunciations or simply casual style – e.g., *iz* for *is*). This is primarily for reasons of consistency and later searchability of the transcript, but also to avoid unwarranted stereotyping of the speaker in question (see Preston 1982 and Bucholtz 2000 for discussions of this point). Where non-standard phonetic forms are relevant to the analysis, they can either be coded using IPA or added in later where relevant.

However, standard orthography is not entirely feasible in unstandardized languages. Auger (pers. comm.) notes that, for her work with Picard, a variety spoken in northern France, the orthography she uses has been developed by the Ch'Lanchron group, who publish books and a quarterly magazine in Picard. It is an analogical orthography, in that it maintains parallels with the orthography of French, a language closely related to Picard and in which all Picard speakers are fluent. However, this orthography is flexible: geographical variants can be spelled differently. For instance, *he was* can be written *il étouot, il étoait, il étot*, depending on how it is pronounced.

Since even in languages as standardized as English, speech often includes "words" that do not have standardized spellings, it is useful to prepare a list of such forms that anyone working on a particular corpus can follow. An example from the Sociolinguistic Archive and Analysis Project (SLAAP) is given in (2):

(2) SLAAP spelling conventions, examples (Kendall 2009)

Uh-huh	Uh-uh	Gonna
Uh-hum	Okay	I'm'a
Mm-mm	Mkay	Wanna
Mm-hm	Nyah	Kinda

Hazen (2010), for a project in which transcriptions are created in a word processor, takes a slightly different approach:[1]

(3) West Virginia Dialect Project (WVDP) spelling conventions (Hazen 2010)[2]

Use the underlying items in the transcript and standard orthography. Do not try to mimic the speaker's speech (i.e. He ain't going to do it. Not – 'e en't gonna do et) . . .

- Type out 'gonna' as 'going to,' 'wanna' as 'want to,' etc. Contractions do not need to be altered.
- Spell numbers (i.e., two thousand and one)
- Don't use abbreviations (i.e. WV) unless the speaker actually says them.
- Do not use ellipsis marks (. . .) because they show up as one character in Word.
- Time Stamps: place every few minutes, or enough that one is visible on the screen at any point of scrolling through the document. More is better! Ex- [12:03]
- Spacing: Single-space the interview but double-space when speakers change.
- Quote marks: insert when needed, including internal dialogue (thoughts).
- Comments: add to the margins using Word's comment feature – not the typescript.
- Transcribe everything that both the interviewer and interviewee say. Never write 'Kirk rambling,' etc.

Examples such as *gonna* and *wanna* indicate how transcription can slip into functioning as coding. Orthographic choices of this type directly affect the use of a transcript for morphosyntactic analysis, as they affect automatic searches. If a corpus is tagged (see Chapter 13), then a formal and explicit level of notation mediates between the representation of speech and the searchable representation of syntactic structures. However, if it is not tagged, the choice of orthographic form is crucial, and systematic notation of any deviation from a standard form (e.g., infinitival *to*) must be noted in the transcription protocol. The only exception, as noted in Section 4, may be when a short transcript is subjected to a one-time analysis, with a focus on qualitative interpretation and no need for searchability or computational tractability.

Linguists differ in decisions regarding the inclusion of non-linguistic sounds (e.g., coughs, laughs, burps), false starts and hesitations, fillers (e.g., *er*, *um*, *y'know*), incomplete (and therefore often uninterpretable) words, and code-switches to a language that is not the focus of investigation. Many of these choices

[1] Samples of transcription and coding protocols mentioned in this chapter are available on the companion website that accompanies this volume (in particular, Valdman 2007; Hazen 2010; Nagy 2011c; Torres Cacoullos 2011).

[2] We note that this excerpt represents a work in progress. The WVDP is now archived in SLAAP. In that version, sixty-seven interview transcripts are time-aligned at the utterance-level (Hazen, pers. comm., October 20, 2012).

are directly determined by the intended use of the transcript and the analytic approach favored by the researcher, discussed next.

A final detail of orthographic representation that the researcher must decide on concerns the imposition of *segmentation* or *punctuation* on spoken language. Speech may be a more or less continuous stream of sound, but it is helpful for readers of a transcript (and possibly for the transcriber as well) to break the stream into segments. Some researchers (e.g., Hazen 2010) work with large chunks, making textual divisions between speaker turns and punctuating sentences. By contrast, Julie Auger uses punctuation insofar as it reflects prosodic organization. She marks pauses, intonational breaks, and interruptions (pers. comm.). Similarly, Rena Torres Cacoullos (pers. comm.) segments transcription into intonation units to provide boundaries that seem more relevant to spoken language.

One danger in the use of punctuation in transcription is different interpretations by users of the transcribed data: a comma may indicate an intonational unit for the transcriber, but may be perceived as marking a pause by a research assistant engaged in coding, and the following phonetic environment may be incorrectly coded as a pause rather than a phonetic segment. For this reason, a detailed and explicit transcription protocol of coding conventions must be used; this is discussed in the final section.

Once again, the particular use of a transcript can determine punctuation choices. When punctuation is used with conventionalized meanings (e.g. upper-case letters for loudness or question marks for rising intonation in conversation analysis), the common preference is to minimize punctuation of any kind other than those transcription conventions. To ease reading of a transcript in such cases, line breaks may be introduced at various natural discourse boundaries. These cases are discussed in the next section. In the case of language documentation, transcribing and linguistic analysis go hand in hand, and it is necessary to revise the form of the transcription repeatedly as the linguist's understanding of the structure of the language develops (see Chapter 4; Jung and Himmelmann 2011: 204).

4 Representing dialogue

Transcripts vary enormously in how faithfully they preserve details of the delivery of talk (i.e., the manner in which speech was produced and the dynamics of the interaction). There is no "correct" level of detail. Indeed, Mishler (1991) has shown how the same interaction has been transcribed differently in research for different analytic purposes. One can argue, however, that there is a correct level of detail for a given research question. As Edwards (2001) notes, the choice of conventions is generally driven by the nature of the interaction and the analytic goal or framework.

As noted earlier, an inevitable trade-off exists between detail in transcription and the amount that can be transcribed. However, feasibility is not the only

consideration. The manner in which an utterance was produced may not be relevant to certain kinds of theoretical analysis, so the goal in transcription is by no means to include as much detail as is feasible. In transcription used for quantitative sociolinguistic analysis, details such as hesitations, overlapping speech, loudness, and other production phenomena are often omitted or simplified. In any kind of analysis, however, the transcriber must always be alert to the potential importance of even these elements. For example, Sharma (2005) noticed that self-repairs in interviewee's echoic usage (structures that paralleled the interviewer's speech) corresponded with certain language ideologies expressed in interviews. This incidental evidence would have been obscured by inexact transcription or omission of either interviewer speech or self-repairs, both of which were initially deemed irrelevant to the core focus of the study.

As we move toward qualitative sociolinguistic modes of analysis, analysis tends to require more faithful documentation of fine details of speech production, interactional structure, and non-verbal activity. Because of the increased attention to these features, transcription for discourse analysis (used broadly here to include discourse, narrative, interaction, and conversation analysis) tends to eliminate the use of any punctuation other than those conventions explicitly listed. To retain readability and to reconstruct the rhythm of the interaction, discourse analytic transcripts use frequent line breaks at boundaries such as turn constructional units (TCUs), intonational phrases, breath groups, or informational phrases (syntactic constituents with a unified intonational contour, often marked by pauses; Gumperz and Berenz 1993). Line numbering is crucial in such transcripts, as are speaker codes.

As part of a wider debate over the principles and practices of conversation analysis and other forms of discourse analysis (see Chapter 21), a fair amount of discussion has taken place over degrees of detail in transcription, with both greater and less detail being critiqued as potentially impeding analysis.

Conversation analysis has developed a particularly detailed set of notation conventions. One common notation system is the Jefferson Notation System (Atkinson and Heritage 1984; Hutchby and Wooffitt 1998; Jefferson 2004). These systems aim to track linguistic and contextual cues in conversation and to model the sequence and timing of an interaction by using notation of the kind illustrated in example (4) (see also Appendix 21.1 in Chapter 21).

(4)

(.)	barely noticeable pause, usually less then 0.2 seconds
(.3), (2.6)	timed pauses
↑word, ↓word	onset of noticeable rise or fall in pitch
A: word [word B: [word	start of overlapping talk (closed brackets '] ' are sometimes used to mark the end of overlap)
.hh, hh	in-breath and out-breath respectively
wo(h)rd	laughter or related style of utterance of word
wor-	sharp termination

wo:rd	lengthening of sound preceding colon(s)
(words)	transcriber uncertain of transcribed words
( )	unclear talk (sometimes each syllable is represented with a dash)
A: word=	
B: =word	no discernible pause between turns
word, WORD	two degrees of increased loudness
°word°	start and end of quieter speech
>word word<	faster speech
<word word>	slower speech
((sniff))	transcriber's notation of non-verbal details

Marginally less detailed conventions that are also widely used include those developed by Gumperz and Berenz (1993), du Bois et al. (1992), and Potter and Hepburn (2005). In many cases, an initial "rough" transcription is used, employing a subset of conventions, and this can subsequently be worked into a much "finer" documentation of talk as action as the researcher's understanding becomes refined through multiple listenings.

In (5)–(7), we illustrate different degrees of detail in the marking of conversational speech. The researcher must decide which of numerous aspects of speech should be represented in a transcript.

In (5), the coding of the transcript reflects much more detail, with particular attention to timing, silence, and breathing. Notice how intuitive characterizations of speech production (e.g., *'lo* in line 3 or *haveta* in line 18) are more acceptable in this context as the data are not being subjected to computerized searches, and are favored to add vivid accuracy to the rhythm of dialogue.

(5) (from Schegloff 2001: 235)

01		1+ rings
02	Marcia:	Hello?
03	Donny:	'lo Marcia,=
04	Marcia:	Yea[:h]
05	Donny:	=[('t's) D]onny.
06	Marcia:	Hi Donny.
07	Donny:	Guess what.hh
08	Marcia:	What.
09	Donny:	hh My ca:r is sta::lled.
10		(0.2)
11	Donny:	('n) I'm up here in the Glen?
12	Marcia:	Oh::.
13		{(0.4)}
14	Donny:	{ hhh }
15	Donny:	A:nd.hh
16		(0.2)
17	Donny:	I don' know if it's possible, but {hhh}/(0.2) } see
18		I haveta open up the ba:nk.hh
19		(0.3)

20	Donny:	a:t uh: (.) in Brentwood?hh=
21	Marcia:	=Yeah:- en I know you want- (.) en I whoa- (.) en I
22		would, but- except I've gotta leave in aybout five
23		min(h)utes. [(hheh)
24	Donny:	[Okay then I gotta call somebody else.
25		right away.
26		(.)
27	Donny:	Okay?=
28	Marcia:	=Okay [Don]
29	Donny:	[Thanks] a lot.=Bye-
30	Marcia:	Bye:.

Schegloff (2001: 236) points out that rendering the above exchange in the approximate format in (6) below would appear to omit little – just silences, breathing, volume, timing – but it is this material, within its sequential context, that indicates the underlying actions being attempted, achieved, and avoided. Note, of course, that the level of detail in the transcript in (6) might be adequate if the focus of the analysis simply dealt with the syntactic structure of requests.

(6) (from Schegloff 2001: 236)
My car is stalled (and I'm up here in the Glen?), and I don't know if it's possible, but, see, I have to open up the bank at uh, in Brentwood?

However, even the detailed transcript in (5) is selective, and by no means exhaustive in terms of transcription detail. If an analysis focuses more on how meaning is conveyed through prosody – specifically, negative evaluation through mimicry in the next example – then the transcriber might choose to include shifts in pitch, as in (7). Even more detail in the transcription of prosodic structure can be achieved by notation systems such as ToBI notation (Tones and Break Indices; Silverman et al. 1992) or interlinear tonetic notation (Cruttenden 1997).

(7) (from Couper-Kuhlen 2001: 24)

The extract is from a phone-in program; M is the moderator of the show and C is a caller

M:	then we go to Hardwick. (.)
	and there we get –
	(.) h sexy Sharon.
	↓hi!
C:	(0.4) °hello° –
M:	{1} °hello° –
	how are you Sharon –
C:	°all right [thanks°
M:	[oh: ↑cheer up dear,
C:	he hh
M:	cheer up;
	for goodness sake;
	don't – don't put me in a bad mood;
	at (.) one o'clock;

Transcripts that are time-aligned with video recording can add further crucial detail of facial or body gesture, direction of gaze, intended addressee, or other contextually disambiguating information, some of which might ultimately be included in the finer transcript presented as part of an analysis (see Chapter 10 for the potential importance of these elements in ethnographic data collection).

In practice, analysts select among available transcription codes of speech production as suits their needs in a particular analysis, and develop new ones for specialized notation (always providing a full list of conventions used). This practice can, consciously or unconsciously, render alternative readings of the data inaccessible to a reader. As in the area of language documentation, therefore, transcription has not been seen as a neutral or mechanical activity in discourse analysis (Ochs 1979; Bucholtz 2000). Indeed, transcription is very much part of analysis in qualitative sociolinguistic research (see Chapter 21 for further examples). In terms of what she calls the *interpretive* dimension (i.e., selecting what material to include in a transcript), Bucholtz (2000) offers an example from a police interrogation that shows how the selective omission of parts of a dialogue as "incomprehensible" produces a very different picture of the motivations of the participants involved. Similarly, in terms of what she terms the *representational* dimension (i.e., orthographic choices), Bucholtz offers an example of how the speech of an African American man is subtly, possibly unconsciously, reshaped in a radio transcript, both standardizing his speech, thereby removing elements of coherence and continuity, as well as retaining random elements for colloquial character. Bucholtz observes that academic transcription is as politically fraught as these instances of "lay" transcripts: whether colloquial detail is retained or omitted, a transcript is always bound to be a representation of an individual's speech that has been heavily mediated by the transcriber/researcher. Ochs (1979) notes that even the choice of column-format transcription (in which each speaker has a different column) or vertical format (in which each speaker follows the previous speaker vertically) might influence the analyst's or the reader's sense of who dominates the interaction. Sensitivity, reflexivity, and transparency in these choices is therefore vital. (See Edwards 2001 for further details on transcribing discourse.)

5 Glossing in multilingual transcription

Variationist sociolinguistic analysis often presents speech data with little markup, highlighting the element under study fairly informally, as in (8).

(8) Sample transcription from a variationist analysis:
I think Ø he thought Ø it was really cool that I spoke French and that I was bilingual. (Liz; Blondeau and Nagy 2008)

This is only possible when the language being studied and the working language are the same, or if the language being studied is well known to the intended

audience. Speech transcribed in a language other than that of the analyst or the published work necessarily involves added layers of *interlinear glossing* and *translation*. As with "monolingual" transcription, speech transcribed in a different language may be used to perform phonological, morphological, syntactic, sociolinguistic, or some other form of analysis; in each case the information that must be included in interlinear notation differs. Even when the language being studied is the same as the working language, interlinear markup can be useful for tracking formal properties of the transcript (this is done extensively in corpus linguistics; see Chapter 13 for details). In this section, we discuss bilingual transcription and in section 6 we briefly review current transcription software that allows bilingual or other types of inter-linear markup.

When transcribing an extract from one language for study in another, formal linguists generally use separate rows to indicate morpheme-by-morpheme glossing and idiomatic translation. In the commonly used Leipzig Glossing Rules (Comrie, Haspelmath, and Bickel 2008), the original language is transcribed on the first row using a regular script or IPA, usually with each word tab-separated. The second row includes translation glosses that align vertically with the relevant word in the first row. These can either consist of whole-word translations, as in (9), or finer standardized abbreviations for morphological detail, as in (10). The final row provides a smooth or idiomatic translation into the working language.

(9) Interlinear glossing using IPA (Faetar)

phonetic (IPA):	u	tʃin	i	awardá	dəvan	də	la	portə
gloss:	the	dog	he	waited	in-front	of	the	door

translation: 'The dog waited in front of the door.' (Speaker F11B; Nagy 2000: 112)

(10) Interlinear glossing using standardized orthography (Beaver Athabaskan)

orthographic:	dáwótt'yedye	aadi
gloss:	what.kind.of.place	3.said

translation: 'She said what kind of place.'
(yaamaadzuyaaze transcr001; Jung and Himmelmann 2011: 209)

Example (10) differs from (9) in using an orthography specific to the language, rather than IPA, and a richer system of morphological description in the gloss. This type of transcription must be accompanied by a glossary of abbreviations, as shown in (11).

(11) Sample of morphological categories used in glosses (Jung and Himmelmann 2011: 209)
Abbreviations: 1,2,3 = first, second, third person (usually indexing the subject argument if not otherwise specified), ANIMO = animated object, ARE = areal, ASP = aspectual, CNJ = conjugation, DIM = diminutive, DU = dual, ELOO = elongated object, HAB = habitual, LOC = locative, FOC = focus, O = object, PFV = perfective, PL = plural, POSS = possessive, PRT = particle, SG = singular, V = valency.

The researcher may choose to make finer distinctions in morphological detail. For instance, the example in (12) distinguishes between cliticization (marked with "="),

suffixation (marked with "–"), and monomorphemic information (marked with "."). Like example (10), this example uses a standardized Romanization rather than IPA in the first line.

(12) Interlinear glossing with morphological detail (Hindi)

radhaa ne hii	bacchon ko	kahaanii	sunaayii
radhaa=ERG=FOC	child–PL=ACC	story.F	hear–CAUS–PERF.F.SG

'It was Radha who told the children a story.' (based on Sharma 2003: 61)

For longer extracts, a reference key of grammatical morphemes can allow the transcriber to focus on simpler lexical glossing and only fill in grammatical detail later, if needed. One concern when developing interlinear glosses for bilingual transcription is the difficulty of dealing with expressions whose semantic value changes across dialects or across different diachronic stages (Ashwini Deo, pers. comm.). In such cases, either a selected semantic variant with variable forms or a selected form with variant meanings must be tracked in the transcripts, possibly with a notation for shifted semantic values across the dataset.

When speakers use more than one language within a single conversation, additional complications exist. As noted above, segments in the "wrong" language may simply not be transcribed. However, when linguists are interested in the full repertoire of speakers, rather than just one of the languages produced, additional markup may be required. Several options are described in Nagy (2012).

6 Transcription software

Current transcription software allows the transcriber to include glossing in a separate *tier*, whereby each entry on the transcript tier is linked to its matching entry on the glossing tier. In such software, translations or glosses are often just one of several tiers of annotation that might be applied to a transcript, whether bilingual or not.

Early (socio)linguistic transcripts were handwritten or typed, and later word-processed, with the end result being a paper document that could be read and marked up. Digital versions have become increasingly searchable and have slowly moved away from traditional text formats. The major shift is to separate content from form in the transcript and to facilitate links between different elements of markup. In this section we describe a few capabilities currently available in transcription software for the basic transcription of data, its coding and annotation, display options, and potential for data sharing.

When selecting software for transcription, it is advisable to consult colleagues and software manuals in order to select the most appropriate and powerful software for the intended use and analysis of the data. We first briefly outline the advantages of various commonly used transcription software (all open-source and available for download at no cost at the time of writing). In general, files can be converted among these different tools.

Transcriber is a graphical user interface tool for speech segmentation and speech transcription. It is used in research involving close phonetic analysis, as its functionalities include spectrograms and energy plots, segmentation of the speech signal and fine manipulation of segment boundaries, and audio playback capabilities. Transcriber specializes in annotation of the speech signal, and allows labeling of speech turns and topic changes; it is not designed for multi-tier morphosyntactic or other annotation, or for fine conversational detail such as overlapping speech. Transcriber is widely used for simple transcription with time codes, rather than for any form of analysis. By contrast, the software packages that we describe next permit transcription as part of diverse linguistic analysis capabilities.

Praat (Boersma and Weenink 2007) is used for transcription with fine-grained time alignment and is also specialized for use by phoneticians (see Chapter 9 for other uses of the software). Advantages for transcription and analysis include automatic annotation, multi-tier phonetic and speaker information, integration with a powerful graphical interface for phonetic analysis, and a scripting facility for specialized automated coding or analysis. Aside from extensive use in phonetics, Praat has also been used in the Sociolinguistic Archive and Analysis Project (SLAAP; detailed instructions and information are available online; see also Kendall 2008, 2009).

CLAN (MacWhinney 2000) is a set of interlinked programs originally developed as part of the CHILDES database for the study of child language acquisition, but now widely used in other fields, such as second language acquisition and sociolinguistics. It currently serves as the standard tool for transcription, coding, and analysis of TalkBank Corpus databases. A transcript can be created and edited in either CHAT (used more in acquisition studies) or CA (used more in conversation analysis) format; these formats can import from and export to other software, such as Praat and ELAN. As is common in such software, standardized formatting for *metadata* encoding is used (i.e., information about the participants and the recording, including any analyst-designated codes; also see discussion in Chapters 4 and 13). CLAN is favored by conversation analysts for several reasons: keyboard shortcuts for classic CA symbols, direct continuous or segmented playback of linked audio/video with highlighting of active segments, and automatic overlap alignment. In addition to these functionalities for transcription, CLAN permits multi-tier annotation of the transcript for specific linguistic analysis (e.g., word class, grammatical information, phonetic features, prosody, or language choice; further details of electronic annotation and markup are covered in Chapter 13). Other advantages of CLAN include compatibility with non-Roman fonts and built-in analysis programs.

ELAN (Wittenburg et al. 2006), produced by the Max Planck Institute of Psycholinguistics, can also be used to annotate audio and video files on multiple linked tiers with time-aligned annotations. ELAN offers more fine-grained, multiple parallelism in annotation than CLAN (e.g., partiture or "musical score" style presentation of multiple speakers), so is well-suited to transcription involving

gestural, postural, and proxemic detail. It has been widely used in the documentation of endangered languages, sign languages, and in sociolinguistics. It includes sophisticated search functions, basic concordance functions, and some statistics regarding frequency of occurrence of different annotated items. Because it is easy to import to and export from, ELAN is compatible with numerous other transcription systems and applications, including Transcriber, CLAN, and Praat. Text is in Unicode (in many different scripts, including IPA) and annotation and transcription files are stored as XML.

This is a small sample of software currently used for transcription. Other software often includes more specialized capabilities. For instance, **TypeCraft**, a web-based system, has the added advantage of permitting multi-party collaboration via a MediaWiki shell, with options for complex tagging, morphological word-level annotation, and an automatic parser. **Fieldworks Language Explorer** (FLEx), produced by SIL, is designed specifically for language documentation and allows for grammatical markup, XML output, morphological analysis and bulk editing, and complex non-Roman script use (see Chapter 4); however, it does not currently have multi-platform or multi-user capabilities. Software is in a constant state of ongoing development and refinement, and we are likely to see advances soon in automated transcription and coding.

Once a transcript is completed, it can be displayed in a number of ways. Kendall (2008: 342) illustrates four different ways of visually presenting transcripts generated from transcription software, including a format much like the traditional text approach, but including time-stamps indicating when each utterance occurs in the recording, and a "graphicalized" version that illustrates the time flow of the conversation but not the text itself. ELAN transcriptions can be exported as traditional text files, but may also be used via ELAN's graphic interface, in which (overlapping) turns of different speakers, the waveform and/or video recording, and tiers for transcription and different types of markup are all simultaneously visible and time-linked. Both ELAN and SLAAP permit links to Praat so that spectrograms or other visual acoustic representations can be displayed and edited, and both also permit playback of any segment of the recording from the same display. These advances allow representation of pauses, overlaps, latching, and other such details, without explicit transcription (though these must still be coded if relevant to analysis). Edwards (2001) illustrates a number of options for displays that arrange speaker turns relative to data codes and/or researcher commentary, including a vertical multi-tier format (the most common choice), column format, or a nested or interspersed format.

A basic text in one file with markup or other annotation in a separate file or separate tiers makes it easier to use the same base transcript for a variety of purposes down the road. Given that the transcription will be marked up and made messy for linguistic analysis, but must be clean and clear for other users, separate files, or separate tiers which can be exported as separate files, are recommended. Creating separate tiers (or separate files) requires distinguishing between the *basic text* and *annotation*, or additional information. In the HLVC

project, we transcribe the speech of the main research participant on one tier and all other speakers, including the interviewer, on other tiers. A new tier, referred to as a *token tier*, is created in which to mark tokens of each dependent variable being examined. This *daughter tier* is linked (time-aligned) to the tier on which the main participant's speech is transcribed. Independent internal (linguistic) variables are coded as daughter tiers to the token tier in which the dependent variable is coded. External (social, stylistic) variables, often spanning longer time segments, may be coded as well. All tokens and codes may then be exported to a spreadsheet or statistical analysis program for quantitative analysis.

Each new version of the transcript – for example, when proofread by a second researcher, or when a new variable is coded – should be stored as a separate file or tier, with a formalized naming convention described in the protocol. This makes it possible to retrace back to the original file if errors or omissions are discovered, or if different practices are applied at different stages of the research project.

Due to space limitations, we do not discuss the *storage* of transcripts in detail here, save to note that the digitization of transcription has led to significant innovations and improvements in this area as well; Kendall (2008) offers a useful discussion of linguistic data storage.

7 Planning transcription: time, transcribers, and accuracy

A common practical question in planning transcription is how much time it is likely to require. The response depends on how much information is to be included in the transcription, whether the transcript is time-aligned with audio/video files, and the level of experience of the transcriber with writing in the language/orthography being used and with transcribing in general. Estimates for native speakers transcribing English orthographically range from 4× (4 hours to transcribe 1 hour of speech) to 10×. In the HLVC project, transcribing rates for different languages being transcribed by research assistants who rarely write their native languages range from 12× to 28×. In this project, Italian, Korean and Ukrainian are transcribed in standard orthography (fastest). Russian is transcribed in standard orthography by keying in Roman characters (transliteration), which are then convertible to Cyrillic via a web-based application. Cantonese and Faetar are transcribed using IPA (slowest).

In Poplack's (1989) Ottawa-Hull French Project, the goal was to maximize the initial rate of transcription, with a follow-up *correction phase*. Two researchers, working with tape recorders and foot pedals (now often replaced by keystrokes to control a digital recording on the same computer as is used for transcribing), after a year of transcription practice, "reached an average transcription rate of a half hour of speech per day" (Poplack 1989: 431) (i.e., a rate of approximately 16×). (We assume these were native speakers of French.) A range of factors are noted that contribute to the range of rates from 7× to 18× (ibid.): congestion of the time-sharing facility or lab, the number of persons participating in the interview, the

rapidity and articulation of their speech, background noise, volume of the recording, and position of the microphone. In a more recent study (Poplack, Walker, and Malcolmson 2006: 194), this time of English, transcription rates are reported in terms of word counts rather than chronological length of the interview, making comparisons difficult. The overall calculation is that the team transcribed 2.8 million words in 2,471 person-hours.

Time-aligned transcription is initially more time-consuming than text-only transcription, but economies are gained in the long term because it is easier to check transcriptions and the broader context via the direct links between the transcription and the recording. Also, broader transcription is feasible as the first pass, with phonetic details being measured or coded later only for relevant segments.

Finally, transcription for qualitative sociolinguistic analysis is naturally far more time-consuming and is therefore frequently limited to carefully selected extracts. In these cases, transcription is an integral part of the analytic process, so requires direct and constant involvement of the primary researcher(s).

This leads to a second common question in planning transcription, namely who should do the transcribing. Once again, this depends on the level of detail to be included, the relative experience of different project participants, and, in many cases, the (non-)availability of funds. As indicated throughout this chapter, there are significant advantages to the researcher doing some or all of the transcription needed, as crucial coding decisions and analytic insights emerge throughout the process. Students and research assistants can certainly be trained and used for some transcription, with both training and financial benefits.

In the HLVC project, both transcription and proofreading are carried out by students who are community members and heritage speakers when possible, otherwise native speakers of a similar variety of the language. Transcribers are generally paid because the work is very slow. However, a number of HLVC transcribers work as volunteers, finding that working with the data is interesting and of potential benefit to their community.

Jung and Himmelmann (2011) highlight the fact that in language documentation work, or indeed any work where the transcriber is not as familiar with the language as the speakers, transcription needs to be conducted in close contact with native speakers, and therefore often in the field (see Chapter 4), with potentially important outcomes:

> working on transcription may lead to the emergence of a new linguistic variety, as it involves the creation of a new written language. This is particularly true in those instances where recorded texts are carefully edited for publication in a local (e.g. educational) context, a process documented, perhaps for the first time, in a rigorous way in Mosel's work on Teop (see Mosel 2004, 2008). But it actually also occurs in similar, though less systematic ways in transcription . . . (Jung and Himmelmann 2011: 202)

Jung and Himmelmann (2011: 205) also note what an unnatural activity transcription is, especially for languages which are not (frequently) written, and make the valuable recommendation that a researcher transcribe a recording

in their own language before engaging in work on another language or in training native speakers to transcribe.

A third central issue in organizing the transcription phase of a project effectively is accommodating the need for corrections and inter-rater reliability checks. Transcriber effects are unavoidable. Anyone who has transcribed recorded data has experienced surprise at discovering that chunks of audio material have been entirely overlooked in the transcript, frequently due to the natural human facility of attending to the salient constituents of a message and tuning out material perceived to be irrelevant to the message. Even for experienced researchers, repeated listening and editing of transcripts is a basic component of producing an accurate transcript. More specific transcriber effects can also arise. Jung and Himmelmann (2011: 208–9) point out that transcribers who are community members may resist transcribing, or transcribing verbatim, certain elements of a recording because of lack of comprehension (possibly due to dialect differences), taboo, disbelief, a desire to tell less or more than is in the recording, and a general (and very natural) concern more for the message than the form of utterances. Sometimes elements are omitted because there is no straightforward translation for them in the linguist's language, as is the case for Beaver evidentials when being translated to English as part of the transcription process (ibid.: 212).

Despite these transcriber effects, some universal and some culture-specific, certain practices can facilitate the accuracy of transcription. Especially in more selective transcription and transcription that is accompanied by coding of the data, it is important that the transcription protocol be well documented. This is vital for replicability by later researchers and also because linguists seem nowhere close to adopting universal standards for transcription, even as they approach it for metadata. Protocols should record decisions such as orthography, punctuation, identification, text formatting, glossing, and tier codes if relevant, and anonymization of speakers and others mentioned. Dated versions should be archived as updates are made so that later researchers can retrieve information accurate to the versions of the transcriptions they use.

Poplack (1989: 433) describes a number of decisions that need to be made regarding ambiguous and non-standard gender and number marking, forms with multiple attested spellings, neologisms, analogical extensions, omissions, additions, and loanwords. She notes that transcribers were encouraged to consult the protocol regularly to ensure consistency in decision-making, at all stages of transcription and correction, and that a simplified version of this protocol would be made available to users of the corpus.

Even with a scrupulous protocol, it remains vital that transcripts be checked several times, by the transcriber as well as, ideally, by another researcher, a stage that can be time-consuming. Poplack (1989: 435–6) calculated that it takes 15–20 hours for a first pass to correct a 2,000-line transcript that had already been passed through an automatic "clean-up" program that fixes recurrent typos, and an additional 10 hours for a second pass. An estimate of one error per 520 words remains

after this careful process, which Poplack suggests is good enough to use the transcription for research without recourse to the recordings. In the Quebec English project, correction is reported to take 2.5–33 hours per interview, for a total of 1,536 hours for three passes over the 2.8 million-word corpus (Poplack, Walker, and Malcolmson 2006: 194–5).

Given the investment of resources for transcription, it is ideal if arrangements can be made for multiple uses of the transcription. The increasing mutual compatibility among transcription and analysis software is allowing linguistics as a discipline to overcome subdisciplinary divides and to share data easily with richer, more robust, and more interdisciplinary results.

References

Atkinson, J. M. and J. Heritage. 1984. *The Structure of Social Action*. Cambridge University Press.

Blondeau, H. and N. Nagy. 2008. Subordinate clause marking in Montreal Anglophone French and English. In M. Meyerhoff and N. Nagy, eds. *Social Lives in Language: Sociolinguistics and Multilingual Speech Communities*. John Benjamins: Amsterdam, 273–314.

Boersma, P. and D. Weenink. 2007. Praat: Doing phonetics by computer. Available at www.praat.org (accessed June 20, 2013).

Bresnan, J., A. Deo, and D. Sharma. 2007. Typology in variation: a probabilistic approach to be and n't in the Survey of English Dialects. *English Language and Linguistics* 11.2: 301–46.

Bucholtz, M. 2000. The politics of transcription. *Journal of Pragmatics* 32: 1439–65.

Childs, B., G. Van Herk, and J. Thorburn. 2011. Safe harbour: ethics and accessibility in sociolinguistic corpus building. *Corpus Linguistics and Linguistic Theory* 7.1: 163–80.

Comrie, B., M. Haspelmath, and B. Bickel. 2008. The Leipzig Glossing Rules: conventions for interlinear morpheme-by-morpheme glosses. Leipzig: Max Planck Institute for Evolutionary Anthropology and University of Leipzig. Available at: www.eva.mpg.de/lingua/pdf/LGR08.02.05.pdf (accessed July 8, 2013).

Couper-Kuhlen, E. 2001. Intonation and discourse: current views from within. In D. Schiffrin, D. Tannen, and H. E. Hamilton, eds. *The Handbook of Discourse Analysis*. Oxford: Blackwell, 13–34.

Cruttenden, A. 1997. *Intonation*, 2nd edn. Cambridge University Press.

Du Bois, J. W., S. Cumming, S. Schuetze-Coburn, and D. Paolino, eds. 1992. Discourse transcription. *Santa Barbara Papers in Linguistics 4*. Department of Linguistics, University of California, Santa Barbara.

Edwards, J. 2001. The transcription of discourse. In D. Tannen, D. Schiffrin, and H. Hamilton, eds. *Handbook of Discourse Analysis*. Oxford: Blackwell, 321–48.

Fox, S. and J. Cheshire. 2011. From sociolinguistic research to English language teaching. *Methods in Dialectology 14*, University of Western Ontario, August 2–6.

Gumperz, J. J. and N. B. Berenz. 1993. Transcribing conversational exchanges. In J. A. Edwards and M. D. Lampert, eds. *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum, 91–121.

Hazen, K. 2010. Step one, step two . . . The big picture of the West Virginia Dialect Project. University of West Virginia ms. (See the companion website for this volume.)
Hutchby, I. and R. Wooffitt. 1998. *Conversation Analysis*. Cambridge: Polity Press.
Jefferson, G. 2004. Glossary of transcript symbols with an introduction. In G. H. Lerner, ed. *Conversation Analysis: Studies from the First Generation*. Amsterdam: John Benjamins, 13–31.
Jung, D. and N. P. Himmelmann. 2011. Retelling data: working on transcription. In G. Haig, N. Nau, S. Schnell, and C. Wegener, eds. *Documenting Languages: Achievements and Perspectives*. Berlin: Mouton De Gruyter, 201–20.
Kendall, T. 2008. On the history and future of sociolinguistic data. *Language and Linguistic Compass* 2.2: 332–51.
2009. *The Sociolinguistic Archive and Analysis Project User Guide. Version 0.96 (second DRAFT–June 2009)*. http://ncslaap.lib.ncsu.edu/pdfs/SLAAP_UserGuide_v096.pdf (accessed July 8, 2013).
Labov, W. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715–62.
2001. *The anatomy of style-shifting*. In P. Eckert and J. Rickford, eds. *Style and Sociolinguistic Variation*, 85–108.
Ladefoged, P. and S. Disner. 2012. *Vowels and Consonants*, 3rd edn. Malden and Oxford: Blackwell.
MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk. Volume I: Transcription Format and Programs. Volume II: The Database*. Mahwah, NJ: Lawrence Erlbaum. Available at: http://psyling.psy.cmu.edu/papers/guides/talkbank.html (accessed July 8, 2013).
Mishler, E. G. 1991. Representing discourse: the rhetoric of transcription. *Journal of Narrative and Life History* 1: 255–80.
Nagy, N. 2000. *Faetar*. Munich: Lincom Europa.
2009. *Heritage Language Variation and Chan*ge. Available at: http://projects.chass.utoronto.ca/ngn/HLVC (accessed 25 June 2013).
2011a. Lexical change and language contact: Faetar in Italy and Canada. *Journal of Sociolinguistics* 15: 366–82.
2011b. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata* XLIII.1–2: 65–84.
2011c. Faetar transcribing guide. University of Toronto ms. (See the companion website for this volume.)
2012. Heritage Language Documentation Corpus (HerLD). Workshop sur l'annotation des corpus multilingues, Paris. Available at: http://projects.chass.utoronto.ca/ngn/pdf/HLVC/Transcribing_multilingual_speech.pdf (accessed 25 June 2013).
Ochs, E. 1979. Transcription as theory. In E. Ochs and B. Schieffelin, eds. *Developmental Pragmatics*. New York: Academic Press, 43–72.
Orton, H., W. J. Halliday, M. V. Barry, P. M. Tilling, and M. F. Wakelin. 1962–71. *Survey of English Dialects: The Basic Material*, 4 vols. Leeds: E. J. Arnold for the University of Leeds.
Poplack, S. 1989. The care and handling of a megacorpus. In R. Fasold and D. Schiffrin, eds. *Language Change and Variation*. Amsterdam: Benjamins, 411–51.
2007. Foreword. In J. Beal, K. Corrigan, and H. Moisl, eds. *Creating and Digitizing Language Corpora*. Houndmills: Palgrave-Macmillan, ix–xiii.

Poplack, S., J. A. Walker, and R. Malcolmson. 2006. An English "like no other"?: language contact and change in Quebec. *Canadian Journal of Linguistics* 51.2/3: 185–213.

Potter, J. and A. Hepburn. 2005. Qualitative interviews in psychology: problems and possibilities. *Qualitative Research in Psychology* 2: 38–55.

Preston, D. 1982. Ritin' fowklower daun 'rong: Folklorists' failures in phonology. *Journal of American Folklore* 95: 304–26.

Sankoff, D. and G. Sankoff. 1973. Sample survey methods and computer-assisted analysis in the study of grammatical variation. In R. Darnell, ed. *Canadian Languages in their Social Context*. Edmonton: Linguistic Research Inc., 7–63.

Sankoff, G., P. Thibault, N. Nagy, H. Blondeau, M.-O. Fonollosa, and L. Gagnon. 1997. Variation and the use of discourse markers in a language contact situation. *Language Variation and Change* 9: 191–218.

Schegloff, E. A. 2001. Discourse as interactional achievement III: the omnirelevance of action. In D. Schiffrin, D. Tannen, and H. E. Hamilton, eds. *The Handbook of Discourse Analysis*. Oxford: Blackwell, 229–49.

Sharma, D. 2003. Discourse clitics and constructive morphology in Hindi. In M. Butt and T. Holloway King, eds. *Nominals: Inside and Out*. Stanford: CSLI Publications.

2005. Dialect stabilization and speaker awareness in non-native varieties of English. *Journal of Sociolinguistics* 9.2: 194–225.

Sharma, D. and L. Sankaran. 2011. Cognitive and social forces in dialect shift: gradual change in London Asian speech. *Language Variation and Change* 23: 1–30.

Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert, and J. Hirschberg. 1992. ToBI: a standard for labeling English prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing*, October 12–16, Banff, Canada, pp. 867–70.

Thieberger, N. and A. Berez. 2012. Linguistic data management. In N. Thieberger, ed. *The Oxford Handbook of Linguistic Fieldwork*. Oxford University Press.

Torres Cacoullos, R. 2011. Transcription notes. Pennsylvania State University ms. (See the companion website for this volume.)

Valdman, A. 2007. NSF Behavioral and Cognitive Sciences grant #0639482.

Wittenburg, P., H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes. 2006. ELAN: a professional framework for multimodality research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*, May 22–28, Genoa, Italy.

Appendix 12.1 Tools and software discussed in this chapter

CLAN	http://childes.talkbank.org/clan
ELAN	Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. http://tla.mpi.nl/tools/tla-tools/elan
FLEx	http://fieldworks.sil.org/flex
Leipzig Glossing Rules	www.eva.mpg.de/lingua/resources/glossing-rules.php
Praat	www.fon.hum.uva.nl/praat
Transcriber	http://trans.sourceforge.net
TYPECRAFT	The Natural Language Database. http://typecraft.org

(All websites accessed July 8, 2013.)