The Routledge Handbook of Language and Identity

Siân Preece

Challenges for language and identity researchers in the collection and transcription of spoken interaction

Publication details
Eva Duran Eppler, Eva Codó
Published online on: 18 Feb 2016

How to cite :- Eva Duran Eppler, Eva Codó. 18 Feb 2016, Challenges for language and identity researchers in the collection and transcription of spoken interaction from: The Routledge Handbook of Language and Identity Routledge
Accessed on: 30 Aug 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Challenges for language and identity researchers in the collection and transcription of spoken interaction

Eva Duran Eppler and Eva Codó

Introduction

The goal of this chapter is to address some of the complexities involved in the process of choosing, collecting and transcribing empirical data to investigate the relationship between language and identity. It provides less experienced researchers working in the field of applied linguistics with the tools for making informed decisions throughout the research process.

The first thing novice researchers need to have clear from the start is what (broad) definition of identity they adhere to, since identity research may be undertaken from a variety of ontological and epistemological perspectives within linguistics and applied linguistics (henceforth AL). The choice of perspective is of fundamental importance because distinct stances demand not only the gathering of distinct pieces of data, but also a different approach to data collection and transcription. For example, social constructionist researchers who follow ethnomethodological principles (see Benwell and Stokoe, this volume) assume that identity categories are emergent in social interaction. This means that they are generally not interested in obtaining background details on informants or the contexts in which the data was collected, since for them what identities are relevant in a given stretch of talk is an emergent, empirically traceable and locally negotiated matter. One type of data, i.e. talk-in-interaction, is usually enough, although it may be gathered in different sites with the purpose of building a collection or comparing across contexts. Researchers try to adopt fly-on-the-wall roles; some even choose to be totally absent from the field to prevent any unwanted disruptions of the naturalness of the focal interactional data. Additionally, since who speakers claim to be is a discoverable matter, transcription procedures search for neutrality. This is reflected in the use of uninformative speaker codes in transcripts, such as AAA or BBB.

By contrast, poststructuralist, interpretivist approaches, such as linguistic ethnography (see Pérez-Milans, this volume), require the researcher to gather different sorts of data, which may of course include naturally occurring talk-in-interaction but not exclusively. Instead, emphasis...
is placed on triangulation and dialogism. Claims about unfolding identity processes are usually made on the basis of ethnographic observations, interviews, recorded social interaction and various other types of data. Researchers are required to self-reflect on how their subjectivities have shaped the research process; they are not mere collectors, but active generators of data (Mason 2009). Transcription practices reflect the situated nature of research: participants are only identified by their real names if they want to (otherwise codes or pseudonyms are used), and their personal trajectories are explained in full detail (see Norton 2013 [2000]). So, what this brief comparison illustrates is that there are a number of decisions that applied linguists doing identity and language research must make very early on in their projects.

This chapter thus addresses some of the issues faced by scholars in the field of applied linguistics and sociolinguistics when collecting and transcribing spoken interaction for the investigation of the relationship between language and identity. The following are questions that researchers embarking on an empirical research project in language and identity research are likely to include:

1. Why did I decide to study identity from a (socio)linguistic perspective?
2. Which framework am I going to work in? What approach/perspective am I going to adopt? (See Part I of this volume.)
3. What are the research questions I intend to answer? (See Parts II–V of this volume.)
4. Which types of data will enable me to answer these questions?
5. Which speakers in which kind of contexts/interactions are likely to provide these data?
6. What’s the best way of collecting these data? Do I do it myself? Do I use a fieldworker? Do I ask the participants to record themselves?
7. How do I prepare the data so that I can analyse it? How do I transcribe it?

Only researchers themselves can answer the first questions. Parts I–V in this volume offer suggestions for them. This chapter mainly addresses questions 4–7 and provides inexperienced researchers with information, guidance and the tools for making informed decisions throughout the process of data collection and transcription.

This chapter is divided into two main sections. The first discusses the most important considerations to be made in choosing and collecting one or several types of spoken data. The second one addresses the principal aspects involved in transcribing the data collected. Each section first provides an overview of the main issues and subsequently reflects critically on how these issues might be addressed. A brief summary of the key points made follows. The chapter closes with a list of related topics and the further reading section.

### Issues in collecting spoken interaction and how to address them: different data types and their critical appraisal

Most research on language and identity in linguistics is based on the examination of spoken interaction, be it casual conversation, institutional talk, interview responses or narrative data. However, it must be borne in mind that powerful insights can also be gained by examining written texts. Sometimes, as with diachronic research, written documents are the only possibility. A case in point is the work by Pavlenko (2004), who analyses the autobiographical memoirs of Russian émigrés to North America at two different historical moments to compare and contrast the ideologies of integration that underpinned these women’s language-learning accounts. Another research strand that draws on textual materials for analysis is the critical discursive perspective (see Zotzmann and O’Regan, this volume). An example is Blommaert
and Verschueren’s (1998) analysis of the construction of national identity in relation to foreign immigration into Belgium by the Belgian progressive press.

Textual and oral materials are not mutually exclusive; they may in fact be conveniently integrated within one single research design. For example, investigating hegemonic (or stereotypical) representations in public discourse of a particular social, ethnic or religious group, whether by mainstream society or by group representatives, is a necessary step to follow in order to understand the practices of self (re)definition of its members, that is, what social categorisations or aspects of it they may be trying to negotiate or simply challenge in their practices/accounts (see Giampapa 2004 for an illuminating case study). Be that as it may, spoken interactive data, whether in combination with other types or on its own, is the most common sort of empirical material used to investigate the relationship between language and identity. This is because it is generally accepted that, on the one hand, identities do not exist independently of language and that, on the other, identity construction is a social matter. In what follows, we shall discuss the most common types of interactional data to be gathered, together with their advantages and limitations.

Naturally occurring interactional data

Various kinds of interactional data may be collected to investigate identities in/through language use. Interactions may involve two or multiple parties; they may occur face to face or be technology mediated; they may be formal or informal; they may have an institutional character or they may not; they may involve social distance and asymmetries of power and knowledge; or they may occur among peers (for a nice collection of studies in a variety of settings, see Antaki and Widdicombe 1998). They may be audio-recorded or video-recorded and used in combination with other types of data or in isolation. As we mentioned earlier, it all depends on the specific questions a researcher wants to answer and the stance s/he takes on what the nature of identity is – where it resides and how it can be best studied. Following common usage, in this section we use the term ‘naturally occurring interactional data’ to refer to unplanned interactive language use that unfolds – insofar as possible – without researcher intervention. This is to distinguish it from (solicited) interview data. Yet many a researcher would rightly claim that interview data is also a naturally occurring type of data, though produced in a specialist type of communicative event.

The advantage of using naturally occurring interactional data is that it enables researchers to examine identity as social action, as part of the business of getting things accomplished in talk. Identity construction in interaction is situated, relational, empirically traceable and, following poststructuralist positions (see Baxter, this volume), multiple, flexible and dynamic. However, it must be borne in mind that the centrality and methodological treatment accorded to interactional data depends, as we said, on researchers’ theoretical point of departure. In fact, what we are dealing with is a continuum of methodological positions rather than discrete, separate choices. From a conversation-analytical perspective, it is fundamental to collect data that is ‘untouched’ by the researcher’s presence. Additionally, conversation analysts depart from the assumption that identity is done in interaction through participants’ ongoing self- and other-categorisation processes (see Benwell and Stokoe, this volume). For that reason, they often make a point of not knowing who the participants are so as not to be ‘biased’ in their analyses by conversationally extraneous identity details. Edwards (1998: 20), for example, claims the following in his analysis of social categories in relationship counselling sessions: ‘I know nothing of any of the three participants (and have never met them) outside of what is said on the
tape.’ By contrast, the researcher’s acquaintance with the researched is not to be refrained from in ethnographic approaches to language and identity, which may include the collection of interactional data. In those studies, it is the researcher’s long-term involvement with informants rather than his/her absence that facilitates participants being oblivious to the researcher’s presence and methods.

As regards the practicalities of data collection, similar considerations must be made when gathering talk for other research purposes (see Cameron 2001). Participants’ informed consent must be obtained and their anonymity guaranteed (see Hultgren et al., this volume). If video-recording is involved, participants and researchers must come to an agreement as to what constitutes good informant protection in ways that abide by university regulations for appropriate research conduct. A major source of difficulty with naturally occurring talk is noise. To ensure good quality recordings, external noise must be kept under control insofar as possible. This may be difficult when recordings take place in institutional settings where a number of social activities may be going on at the same time and participants and other individuals may be manipulating different kinds of artefacts. For instance, recording data in classrooms can be particularly hard going if researchers attempt to capture everything. The best way of going about recording in a classroom is to focus on recording specific types of data at different moments in time (e.g. focused teacher–student interaction, small group work, etc.; cf. Norton and Toohey 2011). Further complexity may be added in professional contexts where interconnected but different courses of action may take place in geographically distant spaces. This is the case in Mondada (2007), where analysing a videotaped operation in a major French hospital involves understanding what goes on in the operation theatre but also in an amphitheatre where the image is transmitted for training and advisory purposes.

**Interview data**

Traditionally, interviews have been considered to be the privileged point of entry into people’s life meanings, identities and experiences. For a long time, it was believed that there was no better way of understanding who people believed themselves to be than to ask them. Yet in recent decades there has been thorough questioning of both interviewing as a method to access people’s values, ideas and emotions, and of identity as having some pre-existing reality status in people’s brains (see Benwell and Stokoe 2006 for a comprehensive critique of this perspective).

Being interviewed is not about telling some inner truth about ourselves; it is a communicative event between two or more interlocutors situated in particular linguistic, socioeconomic, historical and political regimes. This means that interviewees’ responses are shaped by perceptions of the interviewer – that is, who s/he is; what his/her values and expectations might be; what s/he might be after, etc. Interviewers may also be concerned about the larger audience/ readership for their words, and thus construct their responses accordingly. This does not mean that respondents are not being truthful; it just means that they are producing appropriate answers in keeping with their understandings of the interview situation. Of course, interviewees will try to self-present in a positive (or even negative) light, that’s only human, but understanding what constitutes positive or negative self-presentation will provide researchers with valuable information about how they are being positioned as interviewers. This information will then be key for their analytical endeavours. Another source of worry for inexperienced researchers tends to be the veracity of informants’ accounts of the past. It must be taken into account that memories are always reconstructions, which are fashioned and refashioned as they are (re)told. This has nothing to do with informants being sincere. As Block (2006) states, the analytical purchase of
past narratives is not to get to understand informants’ past lives but rather to comprehend their process of self-construction in the present.

The third major pitfall of interviewing as a technique that is still often overlooked is its interactive nature. Responses are shaped by previous questions; they are interactively situated. There is no way in which responses can be rigorously analysed without taking into account the way they were framed by previous interviewer talk, not just the immediately preceding question but the whole (evolving) context of the interaction. With all this, we are not trying to dismiss the interview as a valid research methodology for language and identity research; we just want to make novice researchers aware of its limitations and the ways in which they can be circumvented (for more details on general interviewing procedures see Kvale 2007; Codó 2008).

Talking about one’s identity is no easy task, as it goes to the heart of our most intimate beliefs and emotions. It generates discomfort, and in certain politically charged contexts it may cause stress and even anxiety. For this reason, inquiring directly about identity is generally useless. It will only produce short and ready-made responses. The most suitable approach is to try to engage informants in experiential, self-reflective talk through the formulation of open-ended response prompts (they do not have to be phrased as questions; oftentimes they actually work better if they are not). One ideal field for investigating identity is precisely language (see Preece 2009 as an example of language-based interviews), due to the close (ideological) connection that exists between what language(s) we speak and who we believe we are/want to be. In the case of migration, additionally, local language competence is the symbolic terrain over which issues of legitimacy, moral worth and citizenship are constantly fought.

The semi-structured or unstructured interview format allows researchers the flexibility necessary to formulate context-sensitive questions, pursue emergent topic leads and elicit longish pieces of discourse. One of the elements that researchers will have to consider is the language(s) in which to conduct the interviews, taking into account not only individual repertoires and linguistic competences, but also informants’ will. This may entail having to search for an interviewer who is proficient in informants’ (preferred) language(s) and giving him/her specific training. Researchers will have to decide what to prioritise, taking into account the advantages and disadvantages of each option, whether informants’ voice and ease of expression or their own role as interviewees. Another consideration to make is how many interviews to carry out, how separate in time, and their nature. Of course, the number of interviews will be conditioned by access to informants and practical considerations. Yet this is an important aspect to be thought through before embarking on any research project because it will have an impact on the depth and quality of the data collected.

Among interviewing techniques, perhaps the most popular type for understanding processes of self-definition, self-categorisation, (dis)identification and (dis)affiliation is the life-story interview. Atkinson (1998: 8) defines the life story as a ‘fairly complete narrating of one’s entire experience of life as a whole, highlighting the most important aspects’. One crucial aspect of life stories is that interviewees are in control of the telling. This means that interviewers must deploy a facilitator stance but refrain from trying to steer the narration in particular directions.

Life stories are long interview events, which usually involve (but not necessarily) repeated meetings with informants. The first encounter may be focused on building a relationship of trust, and getting at general details about informants’ past and present lives (and possibly future expectations). Subsequent interviews may be more concrete, purporting either to obtain details of specific events/spheres of informants’ lives (e.g. work, religious practices, social networks, etc.) or the reasoning behind specific life-changing decisions and their significance. Evaluative questions are more risky to ask and more complex to answer than factual questions, and this is
why they tend to be asked later in the research process. One possibility to even out the power asymmetries inherent to any interview situation is to build subsequent encounters around the discussion of transcripts from previous occasions, as in Mills (2004). This is a way of handing control over to informants while simultaneously allowing for the emergence of unexpected foci of interest in interview contexts that may serve to illuminate identity-related issues.

**Narrative data**

Since the 1980s, the popularity of narratives for identity research has continued to grow. This is linked to a major epistemological shift in the social sciences away from abstract, positivist and rational thinking, and towards valuing situated constructions of meaning and emotions. According to De Fina and Georgakopoulou (2012), narrative has flourished because, as a heuristic tool, it encapsulates the ontological and epistemological principles of postmodernism, which values subjectivity (rather than objective reality) and individual experience (rather than factual events).

The use of narrative in identity research follows from biographical approaches to life which hold that becoming social beings is about developing a (coherent) life story about ourselves (Linde 1993), one that is constantly revisited and retold. Narrative is at the heart of the ways in which we construct ourselves as moral individuals, and it is closely associated with de-essentialist understandings of the self, as socially negotiated and relational.

A great deal of research on language and identity has been based on narratives of various sorts: from the life-story narrative mentioned above to specialist narratives of major illness; divorce; religious, ideological or political conversion; career choice, etc. Particularly interesting for applied linguistics are narratives of language conflict (Relaño-Pastor and De Fina 2005). The analytical purchase of narratives in identity research lies in the assumption that it is in narrative that people make sense and give coherence to their fragmented and situated life experiences, and convey a sense of who they are or want to be. Narratives are pervasive in everyday life. For research purposes, we can either tap naturally occurring narratives in professional, family or peer-group contexts, or simply collect research narratives through unstructured or semi-structured sociolinguistic interviews. Although elicited narratives have often been presented as unnatural and constrained, De Fina (2009) advocates a context-sensitive approach to (elicited) narratives which views them as co-constructed and emergent in interaction in the same way as unelicited narratives are.

Atkinson and Delamont (2006) follow a similar line of thinking in cautioning researchers against assuming that narratives are the most ‘authentic route’ to the inner self of informants. As they rightly point out, narratives are forms of self-representation, which are interactively, historically, politically and culturally located. Narratives are not discursive renderings of some pre-existing social reality; instead, it is in narrative that reality is constructed. In Atkinson and Delamont’s (2006: 166) words, narratives have to be analysed in their dual nature as both ‘accounts’ and ‘performances’. As accounts, they are rhetorical devices by which narrators construct their version of past events, legitimise actions and motives, and evaluate others morally; in sum, they are forms of persuasion. Narratives are also situated performances, that is, pieces of discourse produced with and for a given audience, both present and absent. Narratives are thus social events oriented to achieving certain interpersonal effects, such as moving, instilling pity in one’s interlocutors, saving one’s face, etc. Dramatic features play a fundamental role in achieving narrative purpose, and as such they must be central to any analytical endeavour.
Over the last two decades, considerable effort has been made at proposing systematic ways of examining narrative genres, e.g. accounts, and at expanding the notion of narrative beyond Labov’s canonical narrative form. (De Fina and Georgakopoulou 2012 is an excellent overview of major methodological trends and issues, as well as analytical frameworks in narrative inquiry.) There is increasing recognition of narratives as situated events where tellers do identity. Identity is linked to notions such as ‘indexicality’, ‘positioning’ and ‘stance’. Identity work is said to happen at various levels, and in particular at the level of the story-world and that of the interaction. Understanding how identity is done through self-presentation in the telling of particular stories (which are conglomerates of form, content and performance), how this level intersects with the unfolding interactional business of relating to the story recipient (whoever it may be) and how both levels index hegemonic imaginaries of particular social groups has been a major accomplishment of interactional approaches to narrative. Even though these approaches draw extensively on conversationalists’ attention to detail, they also incorporate analytical principles from various other theoretical paradigms (and, accordingly, also other types of data, such as ethnographic details and discursive representations).

To conclude, this section has discussed different types of data to be collected for identity analysis in applied linguistics. As we mentioned at the beginning, they can be used in isolation or in combination. In any case, all the types of data presented can be complemented by participant observation, whose scope and depth will depend on the epistemological paradigm adopted, not only with regard to the nature of identity but also to the relationship between language and social life (see Blommaert and Jie 2010). Deciding on research questions and broad theoretical positions right away is important, because as we have discussed there are important choices to be made as to how much data to be collected and what type(s) from very early on. The following section will introduce researchers to the main issues with regard to data transcription and coding.

**Issues in transcription and how to address them**

Transcription is the process of representing spoken interaction in an (ortho)graphic medium. This includes information on who said what to whom and when, and may include information on how and under what circumstances. The most frequent comments on linguistic transcripts by people who have never seen one before tend to focus on the main features of naturally occurring spoken interaction (and consequently transcriptions thereof): there are repetitions, false starts, retracings (with or without repetition of previously said material), pauses, interruptions and overlaps, to name just a few.

The first question we need to address is whether transcription is still a key issue that researchers who wish to use naturally occurring data for their study of language and identity need to get to grips with. The question poses itself because today the data types discussed in the first section of this chapter tend to be gathered, stored and reproduced digitally. This greatly eases accessing them as a whole or in sections/frames. So why do researchers of spoken interaction still embark on the notoriously time-consuming process of transcription? We do so because even multiple replays do not enable us humans to absorb, store and recall a fraction of the impermanent, synchronous, multidimensional details that constitute interaction. Transcription facilitates data analysis by making interaction available for economical and systematic examination. In the long run, transcription saves time, and this is why analyses are often mainly based on transcripts. The process of capturing real-time interaction in a less transient, written and spatial medium advances our understanding of verbal interaction. As one of the foremost authorities...
on transcription, John Du Bois (1991: 75), puts it, ‘through the experience of transcribing the transcriber is constantly learning about discourse’.

**Transcription is theory**

Although transcription is fundamentally a methodological choice, transcripts are never neutral representations of oral language (Ochs 1979); some researchers even view transcription as a political process (Roberts 1997; Bucholtz 2000).

Transcripts are biased for two main reasons. First, they have to be selective because human interaction is too rich for all of it to be represented (ortho)graphically. If we tried to focus on all aspects of an interactional piece, we could only work on short sections, and it would be virtually impossible to examine one aspect over several interactional episodes. Therefore, researchers/transcribers have to be selective and decide what to transcribe, i.e. what types of information to preserve, which descriptive categories to use, and how to transcribe it (layout and choice of symbols).

Second, decisions made early on in the process of a research project affect later stages. The three main stages of sociolinguistic research projects, adapted from Holmes and Hazen (2013), illustrate that the two activities discussed in this chapter are situated at the later stages.

**Table 19.1** Stages of sociolinguistic research projects

<table>
<thead>
<tr>
<th>Decision stage</th>
<th>Realisation stage</th>
<th>Execution stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectives</td>
<td>Participants</td>
<td>Data collection</td>
</tr>
<tr>
<td>Research questions</td>
<td>Corpora/databases</td>
<td>Transcription</td>
</tr>
<tr>
<td>Theoretical framework</td>
<td>Data collection instruments/method</td>
<td>Coding/tagging</td>
</tr>
<tr>
<td>Units of analysis/linguistic level</td>
<td>method</td>
<td>Accountability/reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analytic method</td>
</tr>
</tbody>
</table>

Adapted from Holmes and Hazen (2013: 181).

The questions identity researchers need to ask themselves (see introduction to this chapter) can now be associated with the three main stages of sociolinguistics research projects. Questions 1–3 ((1) Why did I decide to study identity? (2) Which framework am I going to work in? (3) What are the research questions I intend to answer?) form part of the decision-making process. Questions 4–6 ((4) Which types of data will enable me to answer these questions? (5) Which speakers in which kind of contexts/interactions are likely to provide these data? (6) What’s the best way of collecting these data?) relate to the realisation stage. Question 7 (How do I prepare the data so that I can analyse them?) is a crucial process in the execution of an applied linguistics/sociolinguistic research project.

Data collection and transcription are linked to the project’s objectives, the research questions, and the theoretical approach chosen at the decision stage. Every theoretical framework has a focus that needs to be addressed in transcripts (for the researcher to be able to draw conclusions within the chosen approach); transcription is shaped by theory. Furthermore, every type of spoken interaction data has characteristics that need to be transcribed/coded. For example, only interactions that involve two or multiple parties involve overlaps and lend themselves to a conversation analytical approach. The type of data collected thus also influences – but does not determine – many aspects of the transcription process.
There is no single, unique, correct method of transcription and no unbiased representation of data. The decision about what to transcribe affects not only the researcher who collects the data, but also all researchers who subsequently use the transcripts (especially if the original audio/video recordings are not made available). Decisions on how to transcribe can furthermore affect less experienced researchers’ perception of the data. Elinor Ochs (1979), for example, argues that the speaker represented in the left-most column in a column-based transcription format may be viewed as the principal interactant. We would, however, argue that even fairly novice researchers can see beyond points of data presentation as long as they are aware of the format choices (discussed later in the section on format-based decisions) and reflect on transcription practices. This is what we aim to enable researchers to do with this chapter.

The politics of transcription

Both Celia Roberts and Mary Bucholtz, who work within interactional sociolinguistic and social constructionist approaches respectively, have argued that transcription is a political process. Both of them take up the issue of representing non-standard speech with non-standard orthography (sometimes also called ‘eye dialect’) with the aim of creating a transcript ‘that will look to the eye how it sounds to the ear’ (Schenkein 1978: xi). Apart from the fact that not even phonetic transcripts achieve this, Roberts (1997) points out that the use of non-standard orthography can influence the social evaluation the reader makes of the informants. Bucholtz (2000) demonstrates the inaccuracy and inconsistency of the representation of colloquial speech with non-standard spelling, and refers to the work of Jaffe and Walton (2000: 561), who ‘found that people uncritically and spontaneously read non-standard orthographies as indices of low socioeconomic status’. We question the generality of this statement but agree with Roberts and Bucholtz in calling for a reflexive transcription practice that shows awareness and acknowledgement of the limitations of transcription choices. Non-standard spellings furthermore create considerable problems for computer-readable corpora (the majority today), because of the inconsistency with which lexical items are represented in such transcripts. The word of, for example, has been found to be represented as of,uff, ohv, awv, off, awf and aff in one transcript (Bucholtz 2000). Automatic searches for this word token are only possible if the various tokens of the word are linked to the type in the transcript, with conversion tables, or files which supplement the standard syntax of automated search programmers. Although phonetic or phonological transcription is generally considered to be too time consuming for most types of language interaction research, it may be more practical to transcribe short relevant sections in the IPA (to preserve phonological variation) than use non-standard spellings.

General principles of transcription

The next sections assume that readers who are about to embark on transcription have already made several decisions: they wish to investigate language and identity, they have clear research questions, and they have collected one of the types of interactional data discussed in the first part of this chapter. In the next sections we will therefore mainly discuss transcription issues that emerge in research projects of this type. We will provide researchers with guidelines for choosing/adapting a system of transcription that will allow them to transcribe the features most relevant to their research questions and method of analysis/theoretical framework. What follows draws on the seminal work by Jane Edwards (2003) and John Du Bois (1991).

Most transcripts contain three types of encoding: transcription, coding and mark-up. The more complex and refined the process of transcription becomes, the more it starts to resemble
Challenges for language and identity researchers

coding, the classification of units of data into analytical categories (also referred to as annotation or tagging in corpus linguistics). In language and identity projects, researchers may, for example, want to code passages in which the speaker indexes multiple identities. This example illustrates that coding is even more interpretive and closely tied to theoretical approaches than transcribing. Mark-up is the frequently hidden (behind an interface) and therefore neglected formatting information of transcripts (which also includes metadata/text classifications and structural representations). SGML (Standard Generalised Markup Language) and XML (eXtensible Markup Language) are the most widely used mark-up languages in corpus and computational linguistics. Mark-up is essential for data conversion between different formats/systems and thus crucial for data sharing.

Since the mid-1990s many spoken language projects (MARSEC, HCRC, ToBI, Transcriber, SignStream, CHILDES) have developed the ability to link transcripts to audio and video recordings. This is an important development as it allows researchers to verify transcripts and retrieve information from the original interactional data, which has not been transcribed and/or coded.

When selecting, adapting or developing a transcription system that allows researchers to transcribe the features most relevant to their research questions and theoretical framework, they should ensure that the categories within each contrast set (e.g. short and long pause in the contrast set pause length) of the transcription system satisfy three general principles. More specifically, the categories must be:

- Systematically discernible, i.e. it should be clear whether a section of the recording is a pause or not;
- Exhaustive, i.e. the system should contain a transcription convention or code that allows researchers to transcribe or code each relevant aspect of the data, or allows the addition of categories (plus associated symbols);
- Usefully contrastive, i.e. long and short pauses are defined in relation to the type of data.

These are criteria that every transcription system should fulfil because it is impossible to create an accurate and consistent transcript if one does not know how many and what type of categories there are (in a set).

As virtually all transcripts produced today are machine-readable, the single most important design principle of transcripts is that similar instances are encoded in predictably similar ways. If transcription and encoding is unsystematic, a computer will never find all the instances of what data analysts are looking for; consequently results will be unrepresentative and misleading (because they will contain false positives and/or false negatives). The non-standard orthographic representations of ‘of’ given earlier are an illustrating case in point. Systematic encoding is also important for updating, reformatting and sharing transcripts.

A principle that tends to be particularly important to audiences at the receiving end of research is readability. Readability, however, is not only important for presentation purposes, but also for ‘getting a feel for the data’, hypothesis generation, minimising error in data entry and error checking. This is why many transcription systems, especially those used in spoken interaction research, draw on conventions employed by playwrights (for example, three dots to mark a pause).

Like most written materials, transcripts frequently make use of two main cues for making sense of data: spatial organisation and visual prominence. The use of these two cues is fairly intuitive, but as systematicity is so important for transcripts, it is worth spelling them out. For example, closely related events/types of information tend to be placed spatially near each other;
temporarily earlier events are presented earlier on the page; and information that helps interpret an utterance or event is placed before the utterance. The representation of qualitatively different types of information (e.g. spoken words vs. researcher comments) in distinctly different ways (e.g. different fonts), and the use of few, short, iconic and straightforwardly interpretable symbols (e.g. the use of a slash (/) to mark rising intonation), facilitate readability.

These are the general design principles of transcripts that have emerged in the approximately 22 centuries since spoken language has been rendered in writing (Parkes 1993 cited in Edwards 2003). Naturally, the transcription of video recordings poses even more challenges, as even more information is available than can be captured (ortho)graphically. (For excellent work on multi-modal transcription see Kress and van Leeuwen 2001; Machin 2007; Jewitt 2011.) Going against the general design principles in the transcription of recorded interactional data is comparable to ignoring a red traffic light. Researchers/transcribers may therefore want to save their energy for the decisions they will have to make. These decisions are important, as the options chosen may affect their own and/or others’ perceptions of the interactions presented in the transcript.

**Decisions in transcription**

The decisions researchers working on interactional data have to make fall into two main categories: format-based and content-based decisions.

**Format-based decisions**

Format-based decisions involve the format of the transcripts, and the symbols used to represent features of spoken interaction relevant to the research question. In a column-based format, utterances by different speakers are arranged in separate columns (e.g. in carer–child interactions). The most widely used format in the transcription of naturally occurring interactions is the vertical format, in which utterances by individual speakers are arranged one above the other. This format also helps with the issue of translation and/or inter-linear glossing that many AL scholars who collect data in multilingual contexts face (cf. Creese and Blackledge, this volume; Edwards 2003). Partiture notation (also called ‘musical score’ or ‘stave’ format) is a version of the vertical format which, as the name suggests, resembles a musical score in that utterances/turns by different speakers are put on different lines and the spatial arrangement (frequently in combination with brackets which indicate overlapping stretches of speech) shows when different speakers come in or drop out of the spoken interaction.

**Box 19.1 Partiture notation**

<table>
<thead>
<tr>
<th></th>
<th>Marion:</th>
<th>I see if it can (st-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Carole:</td>
<td>Oh, I have office hours=</td>
</tr>
<tr>
<td>3</td>
<td>Len:</td>
<td>Yeah, that’s ( ) three of em</td>
</tr>
<tr>
<td>4</td>
<td>Rafe:</td>
<td>Is that OK? Could we try that</td>
</tr>
<tr>
<td>5</td>
<td>Carole:</td>
<td>=but I’m here. I mean office hours doesn’t count</td>
</tr>
</tbody>
</table>

Box 19.1, an example of a moderately detailed partiture transcript, which marks overlapping speech (\[\)), truncated words (-), pauses (\(\)), self-completion (=) and rising intonational contour (?), also illustrates that more detailed interspersed transcription or coding can lead to information overload. This is why many researchers nowadays opt for the multi-linear (or inter-linear) format in which codes are placed on separate lines (called tiers) beneath the line that contains the speaker’s utterance (with one-to-one linking between the tiers).

Researchers/transcribers who adopt an existing transcription system will have little choice over the symbols/notations used. How accessible (easily learned), economical, robust and adaptable the symbols used for discourse transcription are may, however, influence researchers/transcribers’ choice of transcription system. The good news about format-based decisions in a digital day and age is that, with consistent encoding and appropriate software, most formats can be easily translated into others.

Content-based decisions

Content-based decisions relate to the types of information that are preserved in a transcript, and the descriptive categories used. Unlike format-based decisions, content-based decisions cannot be adjusted by machine. For every data type and set, and for every research question, a decision has to be made regarding the number and type of descriptive categories used to encode information present in interactional data. Content-based choices have to be matched with those made at the decision stage: the categories not only have to fulfil the general principles of category design outlined earlier (systematically discernible, usefully contrastive and exhaustive), but they also need to be compatible with the theoretical framework adopted, and have to enable the researcher to answer his/her research questions; that is, the categories used need to meet the research objectives. Content-based decisions create distinctions between different transcription systems because they most directly reflect the impact of theory on methods.

The most basic content-based decision involves the units of analysis a researcher wants to investigate (this is why ‘unit of analysis/linguistics level’ is in the first [decision] stage column of Holmes and Hazen’s outline of the research process). In language and identity research, possible choices include a speaker turn, an utterance, an intonation unit or an idea unit. The pros and cons of another important content-based decision, whether to transcribe in standard orthography or eye dialect, has already been discussed. Because language and identity research aims at capturing aspects of interaction as they are perceived by human participants, the categories used in such transcription – just like the transcript itself – are always interpretive. This particularly applies to pause length and prosody. In language and identity research, pause length is considered to be locally negotiated and is therefore more frequently given in terms of interpretive judgements of long and short rather than measured. Similarly it is the researcher’s/transcriber’s task to determine which prosodic cues (prominence/stress, duration/length of sounds/syllables and intonation) are meaningful within a set of data and theoretical framework, and whether and how to encode them. The rhythm of interaction and coordination among participants also require content-based decisions by the researcher/transcriber. The musical score transcript illustrates one way of encoding this type of information. Conversation analysts are particularly interested in turn transition and wish to encode features of naturally occurring interactions such as latching, overlap and interruption. The amount of detail encoded and the number of categories in each contrast set (e.g. self-interruption vs. other interruption or just interruption; beginning and end of overlap or just beginning of overlap) depends on the research question. Researchers who focus on non-verbal aspects and events accompanying spoken interaction still
occasionally encode those aspects that are most relevant to their research questions. Mostly, however, this type of information is nowadays retrieved from time-stamped video recordings that are linked to the transcript. For an illustration and explication of how content- and format-based decisions can lead to quite different transcripts of the same recording by the same researcher (Deborah Tannen), see Mishler (1991: 262–264).

Practicalities

Practicalities are always more appealing than decisions. Researchers who have never transcribed naturally occurring spoken interaction might find it useful to transcribe two to five minutes from their own data before they read on. This will flag up the importance of the general design principles and choices they will be facing when creating representations in writing of the speech events they (or their fieldworkers) recorded. Many practicalities impact on the transcription process: they range from the quality of the recording and the availability of equipment (computers, transcription software, controllers, headsets, etc.), time and space, to the health of the transcriber’s back. We will therefore only include a few that seem particularly relevant to us and would like to advise novice transcribers to have a chat with more experienced researchers (on transcription fora such as http://forum.linguisticteam.org or in person) before they start and whenever they encounter practical problems.

The piece of information that those researchers unacquainted with transcription would probably most like to know about is: how time-consuming is transcription? The answer to this question depends on the quality of the recording, the type of interaction (one, two or multiple participants, frequency of overlaps, etc.) and the amount of detail encoded. Of course, the better the quality of the recording, the easier the transcription process will be. In our experience transcription time lies somewhere between 16:1 and 20:1 (16 to 20 hours to transcribe one hour of recorded interaction), that is, without time stamping and checking.

New transcription and encoding projects and associated software reaches the market virtually every year. If you would like to link your work to a bigger project or contribute to a database (this is a question of data storage and sharing), you may want to use transcription software that facilitates this integration. Most of this software is freely available on the web, so all you need to do is search for CAVA, EAGLES, ELAN, HIAT-DOS, Media Tagger, TED, TEI and the CA format in CLAN, to mention just a few. If you want to integrate your transcripts into a bigger project or database and/or use automated analysis tools, you will have to check the syntax of your files and correct it for the file to become readable by automatic analysis programs.

This leads us back to the single most important design principle of modern – i.e. machine-readable – transcripts: they need to be accurate, consistent, reliable and systematic. To avoid transcriber effects (such as category drift; this is when a feature of spoken interaction first gets transcribed/coded as one category but then slips into a different category during the transcription process) and achieve a ‘Gold Standard’ in transcription, you may want to keep a transcription/coding protocol that records, for example, information on the anonymisation of speakers and the decisions discussed earlier and have the transcript checked by the transcriber after a time interval or, ideally, by another researcher (called inter-rater reliability checks). To achieve the ‘Gold Standard’, this should involve approximately 10 per cent of randomly selected data (Rietved and van Hout 1993).

Transcription is an open-ended process that can start with a ‘rough’ transcript that gets more and more refined until the format and categories used match the research objectives and questions.
Summary
This chapter has examined the main issues faced by scholars in the field of applied linguistics and sociolinguistics when collecting and transcribing spoken interaction for language and identity projects. In summary, when collecting data, the following principles should be borne in mind:

• Before starting, it is important to have a clear definition of identity and of the role of language in its constitution. This will shape research questions, and determine choice of data and methods of data collection/generation.
• Spoken interactional data, be it naturally occurring talk-in-interaction, interviews or narratives, is always contextually situated and interactively constructed. This general principle should always inform data collection decisions. In particular, researchers should be reflexive about their own impact on the production of interview responses and solicited narratives.

The most important points about transcribing, the process of representing spoken interaction in an (ortho)graphic medium, are:

• Like the data collected, transcripts must serve the aims of the research project; that is, they must ‘fit’ the theoretical approach and answer the research questions.
• Transcripts are always interpretive.
• Existing transcription systems/coding schemes should be used and, if necessary, adapted to the specific needs of the research project.
• The transcription should be as accurate, consistent and systematic as possible.

Related topics
Identity in variationist sociolinguistics; Ethnomethodological and conversation analytic approaches to identity; Language and identity in linguistic ethnography; Discursive psychology and the production of identity in language practices; Ethics in language and identities research; The politics of researcher identities: opportunities and challenges in identities research; Constructing age identity: the case of Mexican EFL learners; Being a language teacher in the content classroom: teacher identity and content and language integrated learning (CLIL); Disability identities and category work in institutional practices: the case of a ‘typical ADHD girl’; ‘Comes with the territory’: expert–novice and insider–outsider identities in police interviews; The future of identity research: impact and new developments in sociolinguistics.

Further reading
Blommaert, J. and Jie, D. (2010). *Ethnographic fieldwork: a beginner’s guide*. Bristol: Multilingual Matters. (This introductory book provides very concrete and practical advice on how to conduct an ethnographically designed research project. Using the authors’ examples, it serves to answer most of the questions novice researchers may have. A recommended read before getting into the field.)
Eva Duran Eppler and Eva Codó


Du Bois, J.W. (1991). ‘Transcription design principles for spoken language research’, *Pragmatics*, 1(1): 71–106. (This publication is old but encourages everybody, including upper-level undergraduates, postgraduates and academics, to systematically think through what they want from a transcription and thus helps in choosing the most appropriate system to start with.)

Edwards, J.A. (2003). The transcription of discourse, in D. Schiffrin, D. Tannen and H.E. Hamilton (eds) *The handbook of discourse analysis*. London: Blackwell, pp. 321–348. (This chapter is the most recent, easily accessible publication on the topic by one of the leading international experts, with the added advantage that it focuses on the transcription of discourse.)

References


Challenges for language and identity researchers


