

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

The transcription of interpreting data

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

NIEMANTS, N.S.A. (2012). The transcription of interpreting data. INTERPRETING, 14(2), 165-191 [10.1075/intp.14.2.03nie].

Availability:

This version is available at: <https://hdl.handle.net/11585/658804> since: 2019-01-30

Published:

DOI: <http://doi.org/10.1075/intp.14.2.03nie>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Niemants, N., "The transcription of interpreting data." *Interpreting* 14(2) (2012): 165-191.

The final published version is available online at: https://doi.org/10.1075/intp.14.2.03nie
--

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

<i>This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)</i>

<i>When citing, please refer to the published version.</i>

The transcription of interpreting data

Natacha S. A. Niemants
University of Modena and Reggio Emilia

Abstract

The growing interest in studies based on samples of authentic interpreting data goes hand in hand with the need to think of adequate methodologies to make such data available for research. During the 1980s, transcription issues posed a series of theoretical questions which may now be worth revisiting in the light of recent debates in Interpreting Studies (IS), and of new methodological instruments. Building on the premise that nothing is as practical as good theory, this paper aims at informing decisions on transcribing interpreting events. It first reviews some general theoretical issues involved in any transcription of oral data, and then engages with issues that are particular to IS.

Keywords: oral data, speech corpus, timing, machine-readable, multimodality

Introduction

Research based on samples of authentic interpreting data requires transcription of these data in order to turn evanescent speech into an analysable format. Transcribing confronts the transcriber with a dilemma: a successful response to one requirement (transcribe what you hear) easily results in a failed response to another (make your transcript readable), and the transcriber's compromises will necessarily affect his or her conclusions (Bucholz 2000; Green et al. 1997). Being the result of interpretive choices (i.e. what to transcribe) and representational ones (i.e. how to transcribe), transcription is always theory laden. Yet theoretical issues associated with the transcription process have received relatively scant attention:

In empirical publications, researchers reporting data collection and analysis procedures seldom make mention of transcription processes beyond a simple statement that audio- or videotaped data were transcribed. [...] It is as if these researchers, through their neglect in addressing theoretical or methodological transcription issues, simply assume that transcriptions are transparent, directly reflecting in text the "hard reality" of the actual interaction as captured on audio- or videotape. [...] This is a surprising assumption given that this research methodology has arisen, in large part, through the discovery that language itself is not transparent and hence constitutes a rich source of examinable data. (Lapadat & Lindsay 1999: 65)

Work in Interpreting Studies (IS) has rarely questioned but simply re-employed the methods used in other disciplines, and transcription seems to have been mainly discussed in relation to practical issues pertaining to specific projects (Cencini & Aston 2002; Mack 2006; Martin 2009; Falbo 2009).

Building on the premise that nothing is as practical as good theory, this paper aims at providing an update of transcription theory: while section 1 reviews some general theoretical issues involved in any transcription of oral data, section 2 engages with issues that are particular to IS and to the transcription of interpreting events. As well as the propositions

conveyed by participants, we need to understand their step-by-step realisations and their timings. And we need to have access to large quantities of data which can be analysed by computer in order to identify recurrent behavioural patterns – what interpreters and the other participants do together, moment-by-moment, in collectively managed mediated interactions. To establish such patterns, we require written transcriptions of what participants in recorded interactions were heard to be saying and doing.

My purpose here is to problematize the process occurring between data collection and data analysis by which raw audio data are operationalised (Dam 2001), and thereby to inform decisions on transcribing interpreting data.

1. The transcription of oral data

Transcription is a multidisciplinary concern encompassing a number of fields of research and practice. Although this paper will mainly focus on transcription as a research method, it may be worth devoting a few words to the history of this type of *translation*, in its original meaning of “transfer” and “movement” (see the transcription of Mattioli’s lecture in Nasi & Silver 2009: 194), from the oral to the written form.

Literatures around the world, both in the past and in contemporary times, were and are experienced in the oral mode, through live performances involving audiences, emotions and references to the social or religious context. So called “performance literature”, which is still typical of many non-Western societies, has gradually disappeared in the West, surviving only as written transcriptions representing some or all of the original words (Thomas 2005).

In contemporary life, transcription appears to occur whenever speech is reported in writing. One may find it, for example, in medical and court records, or in parliamentary proceedings and students’ notes (O’Connell & Kowal 1999: 105). It may also be used where non-native and hearing-impaired people may benefit from written subtitles to compensate for language deficits. To these different *uses* of transcription correspond as many kinds of *users*: from service suppliers such as court and parliament stenographers, subtitlers, re-speakers, trainers, and interpreters, to service recipients such as lay citizens, students, and the hearing-impaired, and from private companies producing technological devices to teaching and research centres (Mack 2006).

Another frequent manifestation of transcription is in journalism. Here another question springs to mind: whom should we trust when a journalist states that somebody said something, while that same somebody affirms that he or she never said it? In other words, what should a transcript contain in order to be reliable? How much detail? Some answers to these and other questions are provided by the scientific literature.

Significant contributions have come from the field of phonetics, where transcription has a long history (Roach & Arnfield 1995; Galazzi 2002). Within this framework, the International Phonetic Alphabet (IPA) was born, designed to represent those qualities of speech that are distinctive in spoken language: phonemes, intonation, and the separation of words and syllables. The IPA mainly deals with problems concerning phonetics and phonology, and can be used to highlight the actual sounds of words, to distinguish between varieties of language or to note modifications of sounds when they combine with other sounds. While widely used by foreign language teachers and learners, linguists, speech pathologists and therapists, singers, actors, lexicographers, and even interpreters, IPA transcriptions of any length are much rarer than orthographic ones, since the system requires great attention to detail and is not easy for transcribers or readers to master (O’Connell & Kowal 1999: 108).

In pragmatics and discourse studies, research on transcription has derived principally from the analysis of talk in interaction, where Conversation Analysis (CA) has been characterized by great attention to transcription practices and conventions. Gail Jefferson (Sacks et al. 1974) developed a system where specific conventions were adopted to transcribe: (a) time, date, and place of the original recording; (b) participant identifications; (c) words as spoken; (d) sounds as uttered; (e) inaudible or incomprehensible words or syllables; (f) silences; (g) overlapping speech and sounds; (h) prosodic features (how something is said) such as pace, stretching, stress, and volume (Psathas & Anderson 1990: 79; see Atkinson & Heritage 1984 and Button & Lee 1987 for fuller coverage). The Jeffersonian Transcription System (hereafter JTS), which was born in the pre-digital era, has probably been the most widely used in orthographic (i.e. non-phonetic) transcriptions over the last 30 years, and has paved the way for a number of alternative transcription conventions (see 1.3 below).

With the advent of computers, technology has offered new features to assist the user in effectively turning audio into a written format, and tape cassette and typewriter have been abandoned in favour of transcription machines featuring foot switches, variable speed control, and time-tracking mechanisms, along with digital recording and transcription technology. Machine-readability has become a matter of primary concern, so that transcripts are now required to be both user and computer-friendly. One consequence is that JTS, which was specifically shaped to the typewritten page, has proved insufficient insofar as some of its conventions (in particular those for the transcription of simultaneous speech) are not readily machine-readable.

Ex. 1: Jeffersonian transcription

```
A:      Hello! [How are you?
B:           [Hi           Fine           [thanks.
A:                                           [Good.
```

As can be seen from this example, in the case of an overlap between A and B, one has to use a fixed-space font like Courier (and a fixed maximum line length) in order to align simultaneously spoken parts of utterances vertically on successive lines as JTS does. If one changes the font or the line length, the positions of the overlapping portions will change. The human reader sees what overlaps with what from their relative positions on different lines, but this information is not generally available to computer programmes, for which printed appearance is a secondary issue. Thus there is no way for the computer to know that B's "Hi" overlaps with A's "How are you?" and not with A's next "Good", unless this information is specifically stated in some way. The advent of computer technology has therefore introduced new needs and concepts, like the need for systematic encoding, that is the avoidance of non-meaningful variability, and for *mark-up*, to indicate a system of text annotation.

As alluring as these technological innovations may be – and they do offer a number of possibilities for researchers and transcribers – one potential pitfall of relying on software is the tendency to treat them as “transparent or unmediated” (Rosenthal 2009: 611), and to forget that human beings still play the major role “for the selective input to a computer” (O’Connell & Kowal 1999: 117).

It is not just the transcription product that is important, but also the process, and the transcriber performing it. Researchers can no longer pretend not to see the mediation process involved in this special translation from the oral to the written form, and the role played by the ever wider community of transcribers. Just as translations with parallel text lay bare the transformations, and the losses, involved in the translation process, alignment of audio and transcript uncovers the transformations, and the losses, perpetrated in the transcription

process. Just like translators, transcribers must recognise that they will not be able to make an exact copy of the original (Mack 2006), insofar as “no transcription is a complete record of a spoken event” (Cencini & Aston 2002: 47).

Sections 1.1-1.4 review some general theoretical issues concerning transcription choices which often seem made out of convenience (speed, simplicity, fashionability, etc.), but also need to be born of reflection (is this the best possible solution to my needs?).

1.1 *To interpret*

An extensive phase between data collection and data analysis involves data preparation. This is based on an interpretive process, where selection appears to be the first issue involved.

A transcript that is too detailed is difficult to follow and assess. A more useful transcript is a more selective one. Selectivity, then, is to be encouraged. But selectivity should not be random and implicit. Rather, the transcriber should be conscious of the filtering process. (Ochs 1979: 44)

Selection entails exclusion: just like a good map maker, the transcriber should determine “what to miss out rather than what to include” (Cook 1995: 45), and criteria for exclusion tend to reflect the purposes and ideology of the research.

The dilemma of selection has often been described as opposing objectivity vs. subjectivity (Green et al. 1997), or breadth vs. depth (Thompson 2005). We can also describe it in terms of another dichotomy: *local interpretation* vs. *global sharing*. By local interpretation I mean that each transcription aims at achieving the specific goals of a researcher working within a particular theoretical framework. At the same time, the product of the transcribing activity should be useful for researchers working with other analytical lenses (Leanza 2005: 179).

This conflict has been intensified by the advent of multi-million-word corpora, where more than one researcher may transcribe interactions for both individual and collective purposes. Whereas not long ago researchers used to transcribe tapes they had recorded themselves, nowadays researchers also use transcriptions made by colleagues, students and assistants (see Falbo 2005). One may wonder how far a theoretical apparatus is shared by the transcribers, and whether this makes a difference in the transcription process: transcribers, often unconsciously, already make a first selection when transcribing, and then the researcher, often unaware of what has been selected by the transcribers, makes a further selection of the features he or she is interested in.

Many of these problems are echoed by current authors on transcription as methodology. Thus Zanettin reminds us that the events a researcher may record are one thing; the recordings (audio or video) another; and transcriptions of the latter yet another (Zanettin 2009: 327). Just as recording selects elements from the observable event, a transcription further selects elements from the recording.

For the sake of awareness, selections can be broadly related to six issues every transcriber should ponder during the transcription process: participants, conversational structure, linguistic and paralinguistic features, prosodic features, silences that can or cannot be attributed to participants, and (when dealing with video-recorded data) kinesic elements.

As far as participants are concerned, the ICOR-Conventions developed at the University of Lyon (Groupe ICOR 2007) distinguish between five different types of “identities”, providing suggestions as to how to mark them in the transcripts. Participants can be identified or non-identified; uncertainty of attribution may concern a single participant, or

hesitation between two of them; and there may be cases where the transcriber selects a phenomenon without attributing it to any participant.

The main element of conversational structure is the turn-at-talk. Turns may be produced one after the other, there may be latching between the end of a first turn and the beginning of a second, or there may be overlaps between turns produced by two or more participants at the same time.

Linguistic and paralinguistic features concern “what is uttered”, i.e. the words pronounced by different participants (including cases where they are inaudible or when the transcriber has difficulty in deciding between alternatives) as well as laughter, coughs, inhalations, exhalations, and the like. Prosodic features, on the other hand, describe “how it is uttered”, and cover things like tempo (faster/slower), volume (louder/softer), pitch (higher/lower), duration (lengthening/truncation), and intonation (rising/falling).

The duration of silences can easily be measured using available transcription tools. However their attribution is more problematic issue, since “there are cases where the ‘ownership’ of a spate of silence is in question” (Psathas & Anderson 1990: 89). Simply attributing silence to the following speaker, as many transcribers do, may be in contrast with one of the main principles of conversational analysis, namely that silences, like utterances, are collaboratively managed and negotiated by participants in the interaction (Orletti & Testa 1991: 261). But attribution is clearly highly interpretive, and there may be cases where the transcriber prefers not to attribute silences to specific participants.

The last phenomenon to be affected by the interpretive process falls under the label “multimodality”, which subsumes kinesic elements such as gaze, gesture and body movements. Theories and practices on how to consistently select these kinds of features are currently in the making (Magno Caldognetto et al. 2003; Rohlfing et al. 2006; Groupe ICOR 2007), and generally connected to the second type of choice transcribers need to make, namely how to represent what was selected.

1.2 *To represent*

There are many ways of putting words into writing, making visible and readable what was originally only audible (Falbo 2005: 25). The first decision is whether to use a phonetic or an orthographic transcription (see Thompson 2005 on types and combinations), which largely depends on the purposes of the research. Next, one has to decide what should constitute the basic units and the level of contextual information to be included. Units may be turns at talk or exchanges of turns, clauses or sentence-like objects, tone groups or paratones, words or syllables. Next, how are non-verbal and indeed non-vocal features, such as laughter and applause – or indeed pauses (silent or filled) – to be represented and distinguished from speech? How are interruptions, overlaps, inaudible segments, and dialectal characteristics to be represented? These decisions too depend on research purposes, and are particularly linked to the choice of spatial arrangement for data display: vertical (Ex. 2), column (Ex.3) or partiture (Ex. 4).

Ex. 2: Vertical

A: Did you just get[back]?
B: [Yes], or rather 2 hours ago. It was a great film.
A: Really?

Ex. 3: Column

<u>Speaker A</u> Did you just get [back]? Really?	<u>Speaker B</u> [Yes] rather 2 hours ago. It was a great film.
---	---

Ex. 4: Partiture

A: Did you just get [back]? B:	[Yes], or rather 2 hours ago. It was a great film.	Really?
-----------------------------------	--	---------

(Edwards & Lampert 1993: 11)

Each of these arrangements has its own strengths and weaknesses. The choices made by transcribers enact the theories they hold, constraining the interpretations they, and others, can draw from the data, which in these examples may be influenced by a left to right or top to bottom bias.

The computer era has also brought about new methods to make the transcripts both user-friendly and machine-readable, requiring further reflection on how to represent oral data.

If we want our transcriptions to be machine-friendly, so that a computer can understand the information implicitly contained in a document, we need to make this information explicit for the machine, distinguishing metatextual from textual information and rendering both types non-ambiguous. (Cencini & Aston 2002: 51)

Given that an exhaustive system to suit all purposes may not be achievable, I shall divide attempts into two groups, according to the way in which the data are treated: either on a time base or in terms of hierarchical structuring.

Time-based data models “take the temporal relation between elements as the main principle for the organization of transcription data” (Schmidt 2005: 1). If you look at Figure 1, the words “how are you” are uttered by the doctor at point 1 on the time-line, whereas the name “Ken” and the pronoun “I” are both uttered at point 2, by the doctor and the patient respectively. Overlapping talk by the two speakers is hence visually displayed in what is alternatively known as a “stave” or “single timeline, multiple tiers” data model, where simultaneous events are placed at the same horizontal position, and the left-to-right direction corresponds to temporal sequence (Schmidt 2004: 1).

Time-Line	1	2	3	4	5
Doctor	How are you	Ken			
Patient		I	have a terrible	headache	
...					

Figure 1. Time-based model

Such time-based models enable the transcriber to represent the words uttered by different participating speakers on different tiers, to display overlaps, and to add other tiers to code verbal and non verbal phenomena that are associated with a particular point in time (see Schmidt 2004 for a detailed account of structural relations in this type of data model, and for an overview of transcription tools based on such a stave notation).

In hierarchical data models, on the other hand, “the principal relation between any two elements [...] is not defined by their respective positions on a timeline, but by their positions in an ordered hierarchy” (Schmidt 2005: 2). In Figure 2, “medical visit” is at the top of the

hierarchy and contains two subdivisions. One of those, named “history taking”, is further divided into utterances by the doctor, such as “how are you Ken”, and utterances by the patient, like “I have a terrible headache”. Since there is no link to a time-line in this representation, the overlap between “Ken” and “I” must be explicitly signaled, for instance by adding a * meaning “beginning of overlap” and a ° meaning “end of overlap” in each utterance.

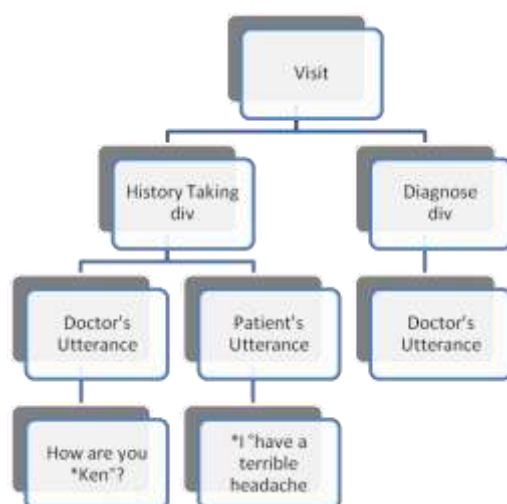


Figure 2. Hierarchical model

Transcription methods using HTML and XML technology can be generally defined as hierarchical, since both the HyperText Mark-up Language (HTML) and the eXtensible Mark-up Language (XML) hierarchically encode electronic documents in order to ease their interchange over the internet (see the TEI guidelines for a detailed account of structural relations in this type of data model: <http://www.tei-c.org/Guidelines>). While working perfectly with written texts, where chapters contain sections, sections contain subsections, subsections contain paragraphs, paragraphs contain sentences and so on, such a hierarchy is often difficult to identify in spoken data, where subdivisions in the interaction may appear subjective (i.e. one might argue that the question “how are you Ken”, where “Ken” is in overlap with the pronoun “I” in the patient’s utterance, is not part of the subdivision “history taking” but of a subdivision one might call “initial greetings”), and where people often seem to “talk in a string” (Tilley 2003: 758).

By and large, both time-based and hierarchical data models try and transcribe such strings, but they do so in different ways. In the former, entities contained in the different tiers of the stave notation can be thought of in reference to a common time-line: simultaneous events are placed at the same horizontal position, and the left-to-right direction within a tier or across tiers corresponds to the temporal sequence. In the latter, entities are assigned to different divisions in reference to a general hierarchy which imposes a more complex abstract structure: following interpretation of their functions, entities are assigned to a particular division, and the order of appearance at any one level corresponds to their temporal sequence.

Time-based data models have the huge advantage of putting things in relation to time, without necessarily requiring them to be distinguished according to their role on multiple levels. As such, they may be suitable to carry out analyses of interactions which basically try to answer the question “Why that *now*?” (Heritage & Clayman 2010: 17, my emphasis). While hierarchical models run the risk of over-interpreting or artificially constructing conversational structure, they also offer an advantage: by relating things to one another, and

not simply to the time-line, they can shed light on particular patterns and practices in terms of the wider context (e.g. question-answer sequences in the context of history-taking).

In principle, the strengths of the two approaches are complementary, and combining them should result in an enhanced view of spoken interaction. In practice, there have been few attempts to bring time-based and hierarchical systems together. Only in 2005 was a scenario concretely proposed where “one particular time-based data model is brought into accordance with one particular subset of the TEI guidelines for transcription of speech” (Schmidt 2005: 3). But given the different theoretical assumptions on which the two models are based, it is still difficult to do in practice what would seem clearly desirable in principle, and conversions between one representation and the other often remain problematic (for a technical account see <http://www.dailyinterpreter.com/archives/1880>).

Most transcription is positioned on a continuum between the needs of local interpretation and those of global sharing and usability, and each interpretative and representative choice is a matter of compromise in this respect. But it also takes place between the opposing instances of *audibility* and *readability*. Those in favour of the written approach will be more bound to the *scripta*, and will hence do their utmost to display visually what they can only hear. Whereas those who promote the oral return will be more bound to the *verba*, and will go back to the oral data to check for features they have not written down. The following section aims at presenting the major concerns of researchers who seem particularly close to one of these two poles.

1.3 *To write or not to write?*

Why should we bother writing down oral data considering that, whatever the choices, “toute fidélité à l’oral n’est qu’illusion” (Galazzi 2002: 142)?

Those with a positivistic world view would answer by arguing that “talk is an observable behaviour” and that “the researcher’s task is to write it down completely and accurately” (Lapadat 2000: 207). In their view, transcribing is worth the effort since the verbatim transcript is “a one-to-one match with the spoken words” and the spoken words are “the sum of the observable event” (ibid.).

Those in favour of interpretivism would instead argue that there is no one-to-one match, since neither speech nor transcripts are transparent. They would also reply, however, that transcribing is worth doing, since “transcription is not simply a way for a researcher to capture, represent, or ‘re-present’ talk, but a constructive and interpretive act in which the researcher positions him/herself” (Lapadat 2000: 209).

Between positivism and interpretivism I would place the search for conventions that led to the Jeffersonian transcription system and then to a number of alternatives, be these the conventions of Du Bois et al. (1993), those developed by Gumperz and Berenz (1993), or those of the more recent VALIBEL (Bachy et al. 2004), ICOR (Groupe ICOR 2007) and PFC (Durand et al. 2009). In all these cases, the reply to the question would be that transcription is worth the effort because it enables researchers to standardize their analyses, and hence to share their results.

Researchers dealing with talk in interaction, such as Gavioli and Mansfield (1990) and Straniero Sergio (1999), among others, would also reply that the process of transcription promotes familiarity with the data, and fosters the methodological and theoretical thinking that is at the basis of data interpretation. As we will see in the next section, this belief substantially matches current approaches to data from interpreting contexts, which are seen as doomed and transient, and hence in need of being written down and resurrected.

Like all speech, interpreting dies on the air. In order to study it, we need to resurrect the corpse by recording and transcribing it, thereby transforming the corpse into a corpus. (Cencini & Aston 2002: 47)

While there may be other reasons underlying the written approach (see Gülich & Mondada 2001), those mentioned may suffice to explain the supremacy of written transcription over alternative methodologies in the Western scientific tradition, which have to do with the different relations that auditory and visual information have to time. To use Chafe's (1995: 54) words, "what is heard is, by its very nature, constantly changing, whereas what is seen stays put long enough to be examined, manipulated, and pondered over".

Practical examples of what can readily be noticed (and hence studied) only in transcriptions provide a further argument in favour of the written approach. This is arguably the case of what one may generically term "emotive communication" or involvement (from Ochs & Schieffelin 1989 to Cirillo 2010): response cries (Goffman 1978), feedback (Schegloff 1982), laughter (Jefferson 1979), and crying (Hepburn 2004).

Writing things down has long seemed the best approach, as it "preserves the data in a more permanent, retrievable, examinable, and flexible manner" (Lapadat & Lindsay 1999: 80). But unsatisfactory results have not prevented researchers from going in the opposite direction, praising the primacy of oral versus written data, and casting doubt on the reliability of observations made on the basis of transcripts. Among the limits they ascribe to the written approach are the difficult standardization process and the huge quantity of data involved.

Despite all efforts at standardization to ease the sharing of transcriptions and results – they argue – the use of a specific notation system for the transcription of interactions is not only impossible but also unscientific, since it "impedes the openness of method required for further creative research and runs the risk of premature certification of current fads" (O'Connell & Kowal 1999: 112). While aiming at a transcript "that will look to the eye how it sounds to the ear" (Schenkein 1978: xi), those who use such conventions also run the risk of forgetting that "a key purpose of transcription is to facilitate the researcher's 'seeing'" (Lapadat 2000: 214), and that

[...] denaturalized transcription, in its faithfulness to oral language, may make speech itself seem alien. This is the paradox of using written texts to represent spoken language. [...] That is, the more a text reflects the orality of speech, the less transparent it becomes for readers unaccustomed to encountering oral features in written discourse. (Bucholtz 2000: 1461)

Even for those accustomed, I would argue, reading written transcripts may sometimes be difficult, since too many conventions have found their own ways of "managing the tension between accuracy, readability, and representation" (Roberts 1997: 170), notwithstanding the use of a full tape-transcribe-code-interpret process that is considered "more complete, accurate, and unbiased for examining language data than alternative approaches, such as online coding or taking field notes" (Lapadat & Lindsay 1999: 80).

Not only is there no transcript that will fit all needs (Lapadat 2000: 215), but there are limits to how much the researcher can reasonably transcribe. One proposed solution is partial transcription on an *ad hoc* basis.

[...] a large database has definite advantages, as the analysis of data in conversation studies usually progresses inductively, the researcher normally does not know at the outset of the research what exactly the phenomena are that he or she is going to focus on. Therefore, it may turn out that she or he wants to analyse events that do not occur

very many times in each single recording [...] Therefore, he or she may need to have access to a relatively large database. In practice, a large portion of the data can be kept as a resource that is used only when the analysis has progressed so far that the phenomena under study have been specified. At that later stage, short sections from the data in reserve can be transcribed, and, thereby, the full variation of the phenomenon can be observed. (Peräkylä 1997: 206)

While those in favour of the written approach counterbalance this limitation by embarking in collective transcription projects and corpus-building, others suggest creating a database of oral data, which can be kept as a resource and then “zoomed-in” on to study specific phenomena which one may eventually transcribe. Since the process of transcription includes analysis at some level, Szakos and Glavitsch (2004) argue that the audio or video recordings – and not the transcripts – are the true data. Their argument is matched by the observation that “speech that is written down is captured and made static and final”, whereas sound recording “can be listened to again and each listening evokes the original event” (Lapadat 2000: 204). Admittedly, “recording per se does not suffice as a guarantee of the reliability of the observations” (Peräkylä 1997: 216). Even a recording will always have to be interpreted. And since neither listening nor transcribing are objective activities, the reliability and validity of observations remain a central concern for any piece of research.

A dialectical stance between the written and the oral poles would seem to best reflect the concerns of transcription makers and users. Once one acknowledges that there is no one-to-one correspondence between the events that unfold during human interaction and what a researcher hears and transcribes in an audio- or video recording, there are two possibilities. Either one selects what to write down and how, remembering that “this selectivity points to a difficulty in developing any one transcript that can be used by different researchers for different purposes” (Lapadat and Lindsay 1999: 73), or one works on the oral, arguing that the search for ever more detailed transcripts is wrong-headed. Irrespective of the approach, be it written with re-listening to the audio, or oral with the transcription of selected extracts, I would argue that one ends up with very similar scenarios: a mixture of written and oral data which raises problematic issues for both the transcriber and the researcher.

I have presented some of the problems raised by the written approach, but the oral one is not without problems either. To begin with, there are cases where the oral data cannot be made publicly available. If one starts from the assumption that research should be replicable, availability of original recordings makes a huge difference. If the recordings are available, and possibly aligned to the transcripts, broad transcription along the lines of subtitles might be the best option, as it could then be made narrower as required for any particular research purpose. But if researchers operate in contexts where there is no way to make the original recording available, their transcriptions should be as detailed as necessary to contextualise the features they are interested in. The privacy law, which regulates the types of information which may be collected and how it may be used and stored, generally depends on four variables: the country where data are collected; the country hosting the server where data are stored; the type of body making the collection, be it private or public; and the purposes of the collection, be they commercial or scientific. Many countries only allow dissemination of written transcripts in which personal data on participants are altered or omitted (see Mondada 2005 on the anonymization of sensitive data). But even when the audio can be publicly accessed, it will still be a long time until computers “will be able to process speech as fast as they handle written data now” (Szakos & Glavitsch 2004: 2). Before that is done, they suggest, “we need some way of linking the speed of text searching with the richness of the speaker’s voice” (ibid.). So while preaching orality, Szakos and Glavitsch also use transcripts, albeit broad ones. We could therefore wonder what would happen if we took their

line of argument a step forward. In other words, how could we practically analyze oral data directly?

Promising developments in this direction have begun to blossom in recent years, so systematic investigation of the problems that may follow in the analysis of interaction should be encouraged. But we should always remember that “the usefulness of a transcription system must be judged in light of the purposes for which it is used”, and that “only the original spoken discourse, some features of which are preserved in an audio (or audio/video) recording, is to be thought of as the data base” (O’Connell & Kowal 1999: 112).

1.4 *To count or not to count?*

The advent of computers and digital devices has had a great impact on transcription methodologies. The digital turn has mainly brought advantages for researchers, who can benefit from small and easy-to-use digital recorders, but it has also shed new light on systems which were suitable for the typewriter but are less so in the computer era. This was the case of the Jeffersonian transcript notation, specifically conceived for the typewriter and for qualitative analyses, but whose dependence upon precise lineation and spacing proved unsuitable for machine-readable purposes and quantitative investigations.

Nowadays virtually all transcriptions of oral interactions are typed and stored on computers. But the moment we speak of large quantities of data, it is advisable to make data not only storable but also analyzable by computer.

Ce saut qualitatif est aujourd’hui indispensable dans la perspective de la linguistique de corpus: il est inutile de stocker de grandes masses de données si celles-ci ne sont pas exploitables par des moteurs de recherche. (Bert et al. 2010: 26)

Leech, Myers and Thomas (1995) describe five stages in the development of “exploitable” speech corpora, which are reduced to four by Thompson (2005): recording, transcription, mark-up and annotation, and application. In short, following the collection and the transcription of oral data, the transcripts shall be computerized in order to be machine-readable. Consequent to this mark-up, the transcriber may wish to add further information, which is referred to as annotation or coding. In the final stage, the corpus can be accessed to count phenomena that have been marked-up and/or annotated.

In corpus linguistics, which is mainly oriented to written texts, quantitative investigations of transcribed speech have generally concerned lexical items that can be counted and localized within utterances, single interactions and/or larger collections. An example of an online corpus of spoken language where sample databases are collected to foster the study of communication is the Talkbank (<http://www.talkbank.org/>).¹

When the original audio and the transcripts are aligned, and the corpora specifically marked-up and annotated, one can also carry out quantitative analyses of phonetic elements. For instance, the PFC project co-developed by the Universities of Ottawa, Toulouse-Le Mirail, Paris X, Oslo and Tromsø provides a phonological overview of contemporary spoken French by counting, *inter alia*, the number of liaisons (linking between words) and schwa (the dropped E) in speakers who differ from a geographic, social or stylistic viewpoint.

One may also work on interactional phenomena such as pauses and overlaps, and quantify their occurrences within speakers’ turns, single interactions and/or larger collections. This is the case in the Belgian VALIBEL databank of oral textual data (<http://www.uclouvain.be/valibel-corpus.html>), and in the French CLAPI data-base of spoken language in interaction (<http://icar.univ-lyon2.fr/projets/corinte/index.htm>). As analyses of the French discourse markers *ben*, *voilà*, *donc* and *alors* clearly show (Bert et al. 2007), the

use of query tools to carry out complex quantitative analyses on transcripts enables one to verify the frequency of certain tokens in relation to others, to the types of interaction and to particular interactional phenomena. For example, one may count the occurrences of “ben” within the turns of a doctor in a doctor-patient interaction, and check how many times it is preceded and/or followed by a pause and/or an overlap. These kinds of complex queries can be extended to cover kinesic elements such as gaze or gesture, as annotation conventions for multimodality are developed (<http://icar.univ-lyon2.fr/projets/corvis/index.html>).

Such quantitative inquiries may be useful to back up qualitative analyses, to test old hypotheses, and to raise new questions, but I agree with Bert et al. that

Les outils fournissent ainsi des « pistes » pour mettre en évidence des associations qu’il convient de vérifier de manière qualitative pour éviter les erreurs d’interprétation, ils ne sont pas à ce jour capables d’identifier de manière automatique des emplois et sont soumis de ce fait à validation. (2010: 26)

Although it is advisable to conceive transcriptions so as to allow for quantitative analyses, it should be remembered that such analyses require human validation. Not differently from transcribing, counting also requires an “intelligent agent” (O’Connell & Kowal 1999: 104) who interprets results as being worth spreading.

Having sketched out some general theoretical issues involved in any transcription of oral data, let us now turn to problems of transcription that are particular to the interpreting context.

2. The transcription of interpreting events

The theoretical issues discussed in the previous section may be summarized as two questions every transcriber should ponder: Can the recordings be made publicly available? And what do I need to capture from the recordings in my transcripts?² While the reply to the first question is often a straightforward “no”, as in the case of interpreter-mediated doctor-patient interactions containing sensitive data that cannot be anonymised, the reply to the second is much longer in the making, and only ends upon finishing transcribing. My purpose is here to problematize some of the issues raised by the transcription of interpreting events, as a way of informing replies to the second question.

The existing literature on interpreting research is not very helpful when it comes to determining an adequate transcription system. Interpreting scholars have only recently turned their attention to transcription as part of research methodology, seeking inspiration in that branch of (socio)linguistics most concerned with the study, hence the transcription, of talk: Conversation Analysis. A basic characteristic of the transcription systems developed within CA is their completeness, as they aim to account for *all* aspects of oral communicative behaviour that might lead to the description of what participants are constructing with their talk-in-interaction. But this completeness makes transcribing an extremely time-consuming activity, and may render the transcripts virtually illegible for non-conversationalists (see above). Researchers within IS have gradually realized that a transcription can be *accurate* without being *complete*, and that there is nothing wrong with a selective transcript:

It is just important that the selection involved in transcribing follows a criterion of *relevance*, that is that the aspects of speech we choose to represent in a (selective) transcript are (all) those that are relevant for the study at hand. (Dam 2001: 171)

Relevance is determined by one's research questions, which in turn determine the features to be analysed, hence the type of transcription that may better help find answers to those questions.

IS has been heavily influenced by the priority traditionally given to research on conference interpreting and monologic modes (simultaneous, consecutive and *chuchotage*), where interpreting is often described in terms of source and target speeches, just as translation is described in terms of source and target texts, and where many corpus-based interpreting studies can be seen as an offshoot of corpus-based translation studies (Shlesinger 1998). This is arguably the case for the European Parliament Interpreting Corpus ([EPIC](#) n.d.), where one can carry out separate searches in transcripts of original texts and/or in transcripts of interpreted texts, for example to compare original English with interpreted English, or English source texts with Italian and Spanish interpreted target texts. Among the research topics one may wish to explore using such monologic interpreting corpora (see Meyer 2008 for another example) are lexical choices, errors, segmentation of input, effects of source language presentation rate, the translation of proper names, and the like (see e.g. Pöchhacker & Shlesinger 2002; Garzone & Viezzi 2002; Garzone et al. 2002 for an overview of possible research questions).

Although most research in IS is conducted within the conference interpreting paradigm, some work has also been done on dialogue interpreting.³ Wadensjö's *Interpreting as interaction* (1998) showed that dialogue interpreters are not only translating source texts into target texts, but also coordinating an interaction in which they actively participate. Thus the kinds of research questions asked of dialogic interpreting corpora (see Meyer 1998; Bühlig & Meyer 2004 for examples) are generally half-way between interpreting and conversational studies, and cover both "traditional" issues such as the interpreters' roles, footings and linguistic choices – where the focus is on the interpreters and their products – and conversational activities such as turn-taking, overlaps and repairing – where the focus is on the co-construction of talk-in-interaction (see Mason 1999 and Gavioli 2009 for an overview of possible research questions).

A happy medium between monologic and dialogic forms of interpreting can be found in TV interpreting corpora like the *Corpus di Interpretazione Televisiva* built at the University of Trieste (Straniero Sergio 2007). Just like any big corpus of spoken language, *CorIT* makes transcription, classification and indexation particularly challenging (Falbo 2009), but is of particular interest in three respects: it covers a number of interpreting modes, thus raising a wide range of research questions; recording interpreters' performances does not influence their behaviour, since recording is an intrinsic feature of this mass medium; and permission for use is not necessary, as TV programmes are readily and publicly available (Cencini & Aston 2002).

I would argue that irrespective of whether their data is monologic or dialogic, transcribers dealing with interpreting face some specific dilemmas concerning genre and purpose. In addition to decisions as to how to interpret and represent spoken language, they need to make further decisions that can be broadly related to participants, languages and timing. Although these choices are strictly interconnected and often influence one another, for the sake of clarity I treat them separately in 2.1-2.3 below.

2.1 *Participants*

Any transcription of speech involving more than one participant needs to indicate the beginning, the end, and the speaker of each utterance. Interpreting data pose a particular problem in this respect, as interpreters have two different participant statuses: they speak both

as the “animators” (Goffman 1981) of talk originally produced by other participants in another language, and as conversationalists in their own right (Cencini & Aston 2002: 49).

When interpreters animate another’s talk, their utterances more or less correspond to one or more utterances produced by a primary participant, which raises the problem of defining the start and the size of the translation unit (Martin 2009; Merlino forthcoming) and of linking the “original” to its “rendition” (Wadensjö 1998). One may then venture as far as differentiating the interpreter’s utterances into close, expanded, reduced and substituting renditions (ibid.), which will require a specific annotation system in order to make these elements subsequently retrievable and analyzable.

One may also wish to classify the way in which participants are presenting their “original” (be it impromptu, read or mixed), and the mode of delivering their “renditions” (be it simultaneous, chuchotage or (short) consecutive). In this last respect, given the terminological confusion that animates IS, where interpreting is variably labeled in terms of modes of delivery (i.e. simultaneous, consecutive), contexts (i.e. conference, public service) and topics dealt with (i.e. medical, business), one needs to make clear-cut distinctions between the type of corpus one is building (monologic, dialogic or mixed), the setting (TV, healthcare services), and the delivery mode (simultaneous or consecutive).

Since participants in an interpreting event matter not only because of what they do but also because of who they are, there may well be an interest in exploring issues of race, class, gender, age, education, nativeness, ideology, etc. which have generally been peripheral in IS. Such a “cultural turn” (Cronin 2002) goes hand in hand with the need to design additional coding conventions to make these elements available for research purposes.

The main interpreting-specific transcription problem regarding participants appears, however, that of segmenting utterances and linking originals with their renditions. This problem is compounded by the fact that there may be more than one original segment corresponding to a single rendition segment, and more than one rendition segment corresponding to a single original one. Furthermore, renditions do not necessarily start immediately after their corresponding original: dialogue interpreters may well first ask for clarifications or perform other coordinating activities (see Baraldi & Gavioli forthcoming for examples). In the light of what we have seen in 1.2, hierarchical data models seem best equipped to cope with these issues, as they enable the transcriber to relate utterances to one another, and not simply to the time-line.

2.2 *Languages*

Another big problem of interpreting data is that they are multilingual. This first implies that whatever tool one decides to use, the transcriber must be able to input language-specific characters. It may also be necessary to annotate each segment or indeed sub-segment with a specification of the language being used – which may not always be self-evident. In addition to this trivial but fundamental issue, particular attention needs to be paid to phenomena that may be present, and hence codified, in any corpus of oral language, but that acquire a special meaning in interpreters’ utterances. This is arguably the case for hesitations and truncations, which could subsequently be treated as indexes of interpreting quality.

Two further interpreting-specific issues are pronoun shifts and code-switches. As for pronouns, in corpus analysis these are generally retrieved either by searching for a specific form (i.e. “I”) or by running queries for this particular part of speech (provided the corpus is POS-tagged). This may be sufficient to observe interpreters’ use of the first vs. the third person, but annotation of pronoun shifts within interpreters’ turns may quickly disclose additional information on the circumstances in which they shift from one to the other (i.e. when they want to make it clear they are speaking for themselves). The same may be said of

code-switches within participants' turns, which may for example open up new questions as to the role of dialogue interpreters in interactions where the primary participants show a certain degree of linguistic competence in each other's language (see Anderson forthcoming; Meyer forthcoming).

Given the overwhelmingly oral nature of interpreting, certain phonetic and prosodic features may also need to be accounted for in the transcription of interpreting data. Transcribing with a phonetic alphabet allows for analyses of participants' pronunciations and accents, thus connecting issues of nativeness or non-nativeness to their actual manifestation in speech. Annotating prosodic features such as tempo, volume, duration, pitch and intonation enables one to compare the prosody of the originals with their renditions, providing evidence for discussions on voice and language quality.

To sum up, language-related transcription issues concern selecting what is uttered, in what language, and how. These selections mainly depend on the purposes of one's research: if the purpose is that of analysing pronoun use, there may well be a temptation to ignore intonation. But "a truly reflexive transcription practice will involve a discussion both of the choices we make and of their limitations" (Bucholtz 2000: 1462). In this specific case one might state, for example, that choosing to disregard prosodic features helps focus on the stated aims of the research, even if it limits the scope of subsequent analyses, and thus the potential of the corpus as a general research (and training) tool.

These kinds of limitations may be counterbalanced by the alignment of transcriptions with digitised audio or video, which seems a *conditio sine qua non* for the analysis of interpreting data. Provided one can align the transcripts with their corresponding audio/video, and possibly different interpretations with the same source speech, choosing a time- or hierarchy-based transcription does not make a real difference.

2.3 Timing

When dealing with interpreting data, it is necessary to pay attention not only to where utterances start and end but also to how they intertwine. In multi-party conversation, participants will occasionally talk at the same time, for instance to co-produce another speaker's utterance, or interrupt in order to signal disagreement or to compete for the floor. In interpreting contexts, however, overlaps do not necessarily have these conversational implications, since two distinct floors of talk may be involved: one in the first language (L1) and one in the second (L2).

Provided that they do not interfere acoustically, overlaps between talk on different floors are not generally treated as significant: in simultaneous or *chuchotage* mode, an interpreter may be able to talk in the L2 without being perceived as interrupting or contributing to the L1 talk which s/he is interpreting. Where the interpreter functions as amplifier (in either simultaneous or consecutive mode), s/he is in fact usually excluded from participating in the first floor, while having a near-monopoly in the second (Cencini & Aston 2002: 50).

Thus, in transcribing interpreting data one may wish to annotate overlaps according to their floor status, distinguishing between conversational overlaps (on one floor) and synchronous speech (on two or more distinct floors).

I would argue that irrespective of floor status, representation in a time-based score format is best equipped to cope with timing issues. In the case of dialogic forms of interpreting, where the interpreter speaks on the same floor as primary speakers, the format immediately shows when one turn follows directly after another has been completed, or

whether it is preceded by pauses and hesitations, thereby giving a clearer picture of “the cooperation between the primary actors and the interpreter” (Meyer 1998). In the case of monologic forms of interpreting, where interpreters speak on a distinct floor, the format shows interpreters’ *décalage*, that is the time they wait before beginning their target language renditions of the unfolding originals. This helps reflect on “le déclenchement et la taille de l’unité d’interprétation” (Martin 2009: 122), and can arguably reveal something about the cognitive dimensions of the interpreting process (Meyer 1998: 80).

Conclusion

I have identified a number of theoretical issues, ranging from the general one of transcription as a special type of translation involving interpretive and representative choices, to more specific ones related to the transcription of particular types of interpreting events. In the belief that nothing is as practical as good theory, I have only touched on a few of the many practical problems linked to specific research purposes and specific transcription tools.

The kinds of issues mentioned in this paper need to be addressed by any transcriber dealing with interpreting data. Since they can neither use technology to bypass transcription nor simply abandon the entire process, transcribers have to be reflective about the procedures they adopt, as should transcript users – as critical readers and analysts. Two further issues alluded to may be briefly noted in conclusion.

Whenever they choose what to include in a transcript, transcribers want to end up with something useful both for themselves and for others. However, the belief that the more we include the better may be mistaken. Selection is advisable in order to maintain one’s focus on the stated aims of a specific project, and the addition of project-specific mark-up and annotation should be kept to a minimum, or else done in such a way as to be easily disregarded, in order not to limit utility for other researchers.

This leads to a second issue, which is the need to present transcripts and connected analyses to non-specialist audiences or to scholars in other fields. Although clearly beyond the scope of this paper, I would argue that this points in favour of multi-layered transcriptions, where different layers serve to add different sets of features and annotations to a basic transcript of the words said, allowing different export options. Be it by using stave notation tools, or by saving a broad transcription before adding other features, a multi-layered approach is also practical from the researcher’s viewpoint, since “it is not unusual for the transcriptionist to return to the same data and re-transcribe more ‘precisely’ such matters as these – because, as we have noted, ‘precision’ is related to the purpose of study” (Psathas & Anderson 1990: 88).

Bucholtz (2000), in his call for greater thought about whether, when, and how one uses transcription, stresses that the main goal should not be objectivity or universality, but responsibility.

Ultimately, what is needed is a reflexive discourse analysis in which the researcher strives not for an unattainable self-effacement but for vigilant self-awareness. (2000: 1461)

Thus researchers in Interpreting Studies need to be mindful of transcription theories and methods, finding their own compromises between comprehensiveness and comprehensibility. From my own search for a happy medium, I draw the three following lessons.

The first is for transcription users:

Transcripts are not transparent or unmediated, [...] they reveal a great deal about the text’s presuppositions, and [...] even without being familiar with the source language

one can still bring a critical eye to how those presuppositions are being deployed. (Rosenthal 2009: 611).

As I have argued above, like translations with parallel texts, transcripts with parallel audio/video can uncover the transformations and the losses perpetrated in the transcription process. This does not mean that Interpreting Studies should necessarily follow the example of work on translation, which has for centuries compared translation(s) with original(s) in order to shed light on the translator's choices. It rather means that following the technical advances that have made audio and video alignment attainable, researchers can no longer pretend not to see the mediation of the "intelligent agent" who selectively turns the oral into a written format.

The second is for transcription analysts:

It is all too easy to take data in the form provided and move on to whatever kind of analysis one favours, being glad not to get one's hands dirty with the processes of collecting and transcribing. But one cannot fully understand data unless one has been in on it from the beginning. (Chafe 1995: 61).

In quantitative analysis based on large corpora, researchers can often do little but accept the transcripts made by others. But in qualitative analysis, analysts should resist the temptation to accept transcriptions as given and be prepared to return to the oral data.

The third is for transcribers – the transcription makers. Because research purposes differ, there can be no universally applicable guidelines for transcription. We might, however, develop guidelines for listing the criteria on which a transcription has been based. This might help to facilitate and improve analyses, as well as to encourage collaboration in creating data banks that can suit both local and general purposes. While not claiming to have in any way formulated such a set of guidelines, I hope to have provided some ideas to ponder when carrying out research which makes use of recordings of interpreting events as its primary data source.

Notes

1. Anyone willing to use the data can become a member by contacting Brian MacWhinney (macw@cmu.edu).

2. This question may be further divided following O'Connell and Kowal (1999: 115-116):

- a) How much should be encoded in a transcription system?
- b) How should these features be encoded?
- c) How should the phenomena encoded for transcription be operationalized?
- d) For whom is a transcript intended?
- e) What, then, is to be said of standardization?

3. In the present paper, the label "dialogue" interpreting is used to subsume liaison, community, public service and a number of other types of interpreting, as it best reflects the dialogic nature of a practice involving at least three parties (see Mack 2005 for a terminological discussion).

References

Anderson, L. (forthcoming). Code-switching and coordination in interpreter-mediated interaction. In C. Baraldi & L. Gavioli (Eds.), *Coordinating participation in dialogue interpreting*. Amsterdam/Philadelphia: John Benjamins.

Atkinson, J. M. & J. Heritage (Eds.) (1984). *Structures of social action*. Cambridge: Cambridge University Press.

Bachy, S., Dister, A., Francard, M., Geron, G., Giroul, V., Hambye, P., Simon, A. C. & Wilmet, R. (2004). Conventions de transcription régissant les corpus de la banque de données VALIBEL. Université Catholique de Louvain.

http://www.uclouvain.be/cps/ucl/doc/valibel/documents/conventions_valibel_2004.PDF (accessed 10 January 2012).

Baraldi, C. & Gavioli, L. (Eds.) (forthcoming). *Coordinating participation in dialogue interpreting*. Amsterdam/Philadelphia: John Benjamins.

Bert, M., Bruxelles, S., Etienne, C., Mondada, L., Plantin, C., Traverso, V. & Valero, D. (2007). L'étude des particules à l'oral dans différents contextes à partir de la banque de données de Corpus de Langue Parlée en Interaction CLAPI. Actes des 5^{èmes} Journées de Linguistique de Corpus JLC5, Lorient, septembre 2007. http://icar.univ-lyon2.fr/projets/corinte/bandeau_droit/ressources.htm (accessed 29 October 2011).

Bert, M., Bruxelles, S., Etienne, C., Jouin-Chardon, E., Lascar, J., Mondada, L., Teston, S. & Traverso, V. (2010). Grands corpus et linguistique outillée pour l'étude du français en interaction (plateforme CLAPI et corpus CIEL). *Pratiques* 147-148, 17-34.

Bührig, K. & Meyer, B. (2004). Ad hoc interpreting and achievement of communicative purposes in briefings for informed consent. In J. House & J. Rehbein (Eds.), *Multilingual communication*. Amsterdam/Philadelphia: John Benjamins, 43-62.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics* 32 (10), 1439-1465.

Button, G. & Lee, J. R. E. (Eds.) (1987). *Talk and social organization*. Philadelphia: Multilingual Matters.

Cencini, M. & Aston, G. (2002). Resurrecting the corp(us|se): Towards an encoding standard for interpreting data. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities*. Amsterdam/Philadelphia: John Benjamins, 47-62.

Chafe, W. (1995). Adequacy, user-friendliness, and practicality in transcribing. In G. Leech, G. Myers & J. Thomas (Eds.), *Spoken English on computer*. New York: Longman, 54-61.

Cirillo, L. (2010). Managing affect in interpreter-mediated institutional talk: Examples from the medical setting. *The Journal of Specialised Translation* 14. http://www.jostrans.org/issue14/art_cirillo.php (accessed 4 May 2012).

Cook, G. (1995). Theoretical issues: Transcribing the untranscribable. In G. Leech, G. Myers & J. Thomas (Eds.), *Spoken English on computer*. New York: Longman, 35-53.

- Cronin, M. (2002). The empire talks back: Orality, heteronomy and the cultural turn in interpreting studies. In F. Pöschhacker & M. Shlesinger (Eds.), *The interpreting studies reader*. London/New York: Routledge, 387-397.
- Dam, H. V. (2001). The manipulation of data: Reflections on data descriptions based on a product-oriented PhD in interpreting. In A. Gile, H. V. Dam, F. Dubsiaff, B. Martinsen & A. Schjoldager (Eds.), *Getting started in interpreting research: Methodological reflections, personal accounts and advice for beginners*. Amsterdam/Philadelphia: John Benjamins, 163-183.
- Du Bois, J. W., Schuetze-Coburn S., Cumming, S. & Paolino, D. (1993). Outline of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum, 45-90.
- Durand, J., Laks, B. & Lyche, C. (Eds.) (2009). *Bulletin PFC N°1*. Projet Phonologie du Français Contemporain. http://www.projet-pfc.net/bulletins-et-colloques/cat_view/918-bulletins-pfc/919-bulletin-pfc-nd1.html?limit=5&limitstart=0&order=name&dir=DESC (accessed 12 January 2012).
- Edwards, J. A. & Lampert, M. D. (Eds.) (1993). *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum.
- EPIC (n.d.). European Parliament Interpreting Corpus. <http://sslmitdev-online.sslmit.unibo.it/corpora/corporaproject.php?path=E.P.I.C.> (accessed 14 March 2008).
- Falbo, C. (2005). La transcription: une tâche paradoxale. *The Interpreters' Newsletter* 13, 25-38.
- Falbo, C. (2009). Un grand corpus d'interprétation: à la recherche d'une stratégie de classification. In *Cahiers de recherche de l'école doctorale en linguistique française*. Milano: Lampi di stampa, 105-120.
- Galazzi, E. (2002). *Le son à l'école*. Brescia: La Scuola.
- Garzone, G. & Viezzi, M. (Eds.) (2002). *Interpreting in the 21st century: Challenges and opportunities*. Amsterdam/Philadelphia: John Benjamins.
- Garzone, G., Mead, P. & Viezzi, M. (Eds.) (2002). *Perspectives on interpreting*. Bologna: CLUEB.
- Gavioli, L. & Mansfield, G. (Eds.) (1990). *The PIXI corpora*. Bologna: Cooperativa Libreria Universitaria Editrice.
- Gavioli, L. (Ed.) (2009). *La mediazione linguistico-culturale: una prospettiva interazionista*. Perugia: Guerra.
- Goffman, E. (1978). Response cries. *Language* 54 (4), 787-815.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

- Green, J., Franquiz, M. & Dixon, C. (1997). The myth of the objective transcript: Transcribing as a situated act. *TESOL Quarterly* 31 (1), 172-176.
- Groupe ICOR (2007). Convention ICOR. http://icar.univ-lyon2.fr/projets/corinte/bandeau_droit/convention_icor.htm (accessed 29 January 2012).
- Gülich, E., & Mondada, L. (2001). Konversationsanalyse, Analyse conversationnelle". In G. Holtus, M. Metzeltin & C. Schmitt (Eds.), *Lexikon der Romanistischen Linguistik*. Tübingen: Max Niemeyer, 196-250.
- Gumperz, J. & Berenz, N. (1993). Transcribing conversational exchanges. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum, 91-122.
- Hepburn, A. (2004). Crying: Notes on description, transcription, and interaction". *Research on Language and Social Interaction* 37 (3), 251-290.
- Heritage, J. & Clayman, S. (2010). *Talk in action*. Chichester: Wiley-Blackwell.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology*. New York: Irvington, 79-96.
- Lapadat, J. C. (2000). Problematizing transcription: Purpose, paradigm and quality. *International Journal of Social Research Methodology* 3 (3), 203-219.
- Lapadat, J. C. & Lindsay, A. C. (1999). Transcription in research and practice: From standardization of technique to interpretive positioning. *Qualitative Inquiry* 5, 64-86.
- Leanza, Y. (2005). Roles of community interpreters in pediatrics as seen by interpreters, physicians and researchers. *Interpreting* 7 (2), 167-192.
- Leech, G., Myers, G. & Thomas, J. (Eds.) (1995). *Spoken English on computer*. New York: Longman.
- Mack, G. (2005). Interpretazione e mediazione: alcune osservazioni terminologiche. In M. Russo & G. Mack (Eds.), *Interpretazione di trattativa : la mediazione linguistico-culturale nel contesto formativo e professionale*. Milano: Hoepli. 3-17.
- Mack, G. (2006). Detto scritto: un fenomeno, tanti nomi. *InTRAlinea* Special issue: *Respeaking*. http://www.intralinea.it/specials/respeaking/ita_more.php?id=464_0_41_0_M (accessed 17 March 2010).
- Magno Caldognetto, E., Cosi, P., Cavicchio, F. & Poggi, I. (2003). La partitura: un sistema di annotazione multilivello di interazioni multimodali basata su ANVIL. *Atti delle XIV Giornate del GFS*, Viterbo, 4-6 dicembre 2003. http://www.istc.cnr.it/doc/72a_2005060715466t_em-GFS2003-03.pdf (accessed 17 March 2010).

Martin, P. (2009). Grand corpus oraux: classification, transcription, exploitation des données. In *Cahiers de recherche de l'école doctorale en linguistique française*. Milano: Lampi di stampa, 121-134.

Mason, I. (Ed.) (1999). *Dialogue interpreting*. Special issue, *The Translator* 5 (2).

Merlino, S. (forthcoming). Traduction orale et organization de la parole. In L. Mondada (Ed.), *Corps en interaction*. Lyon: Editions ENS.

Meyer, B. (1998). What transcriptions of authentic discourse can reveal about interpreting. *Interpreting* 3 (1), 65-83.

Meyer, B. (2008). Interpreting proper names: Different interventions in consecutive and simultaneous interpreting?. *Transkom* 1 (1), 105-122.

Meyer, B (forthcoming). Ad hoc interpreting in multilingual work meetings: Who translates for whom? In C. Baraldi & L. Gavioli (Eds.), *Coordinating participation in dialogue interpreting*. Amsterdam/Philadelphia: John Benjamins.

Mondada, L. (2005). Constitution de corpus de parole-en-interaction et respect de la vie privée des enquêtés: une démarche reflexive. Rapport sur le projet Pour une archive des langues parlées en interaction. Statuts juridiques, formats et standards, représentativité, financé par le Programme Société de l'Information/Archivage et patrimoine documentaire, mars 2005. http://icar.univ-lyon2.fr/projets/corinte/documents/Mondada_juridique_MARS05.pdf (accessed 17 March 2012).

Nasi, F. & Silver, M. (Eds.) (2009). *Per una fenomenologia del tradurre*. Roma: Officina Edizioni.

O'Connell, D. & Kowal, S. (1999). Transcription and the issue of standardization. *Journal of Psycholinguistic Research* 28 (2), 103-120.

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics*. New York: Academic Press, 43-72.

Ochs, E. & Schieffelin, B. B. (1989). Language has a heart. *Text* 9 (1), 7-25.

Orletti, F. & Testa, R. (1991). La trascrizione di un corpus di interlingua: aspetti teorici e metodologici. *Studi Italiani di Linguistica Teorica ed Applicata* 20 (2), 243-283.

Peräkylä, A. (1997). Reliability and validity in research based on transcripts. In D. Silverman (Ed.), *Qualitative research: Theory, methods and practices*. London/Thousand Oaks/New Delhi: Sage, 201-220.

Pöhhacker, F. & Shlesinger, M. (Eds) (2002). *The interpreting studies reader*. London/New York: Routledge.

Psathas, G. & Anderson, T. (1990). The practices of transcription in conversation analysis. *Semiotica* 78 (1/2), 75-99.

Roach, P. & Arnfield, S. (1995). Linking prosodic transcription to the time dimension. In G. Leech, G. Myers & J. Thomas (Eds.), *Spoken English on computer*. New York: Longman, 149-160.

Roberts, C. (1997). The politics of transcription. Transcribing talk: Issues of representation. *TESOL Quarterly* 31 (1), 167-171.

Rohlfing, K., Loehr, D., Duncan, S. Brown, A., Franklin, A., Kimbara, I., Milde, J.-T., Parrill, F., Rose, T., Schmidt, T., Sloetjes, H., Thies, A. & Wellinghoff, S. (2006). Comparison of multimodal annotation tools: Workshop report. *Gesprächsforschung* 7. <http://www.gespraechsforschung-ozs.de/heft2006/tb-rohlfing.pdf> (accessed 29 December 2009).

Rosenthal, A. (2009). Lost in transcription: The problematic of commensurability in academic representations of American Sign Language. *Text and Talk* 29 (5), 595-614.

Sacks, H., Schegloff, E. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696-735.

Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk*. Washington DC: Georgetown University Press, 71-93.

Schenkein, J. (Ed.) (1978). *Studies in the organization of conversational interaction*. New York: Academic Press.

Schmidt, T. (2004). Transcribing and annotating spoken language with EXMARaLDA. In *Proceedings of the LREC-Workshop on XML based richly annotated corpora*, Lisbon 2004. Paris: ELRA. <http://www.exmaralda.org/publikationen.html> (accessed 29 December 2010).

Schmidt, T. (2005). Time-based data models and the Text Encoding Initiative’s guidelines for transcription of speech. *Arbeiten zur Mehrsprachigkeit*, Folge B, 62.

Shlesinger, M. (1998). Corpus-based interpreting studies as an offshot of corpus-based translation studies. *Meta* 43 (4), 486-493.

Straniero Sergio, F. (1999). The interpreter on the (talk) show: Interaction and participation frameworks. *The Translator* 5 (2), 303-326.

Straniero Sergio, F. (2007). *La mediazione linguistica nella conversazione spettacolo*. Trieste: EUT.

Szakos, J. & Glavitsch, U. (2004). Portability, modularity and seamless speech-corpus indexing and retrieval: A new software for documenting (not only) the endangered Formosan Aboriginal languages. *E-MELD Language Digitization Project Conference, Workshop on Linguistic Databases and Best Practice, Detroit, USA, July 15-18, 2004*. <http://emeld.org/workshop/2004/szakos-paper.html> (accessed 23 December 2009).

Tilley, S. A. (2003). Challenging research practice: Turning a critical lens on the work of transcription. *Qualitative Inquiry* 9, 759.

Thomas, R. (2005). Performance literature and the written word: Lost in transcription?. *Oral Tradition* 20 (1), 1-6.

Thompson, P. (2005). Spoken language corpora. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice*. Oxford: Oxbow Books. 59-70.

<http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm> (accessed 23 December 2010).

Zanettin, F. (2009). Corpora multimediali e analisi dell'interazione. In L. Gavioli (Ed.), *La mediazione linguistico-culturale: una prospettiva interazionista*. Perugia: Guerra, 325-355.

Wadensjö, C. (1998). *Interpreting as interaction*. London/New York: Longman.

Author's address

Natacha Niemants
Via A. Spinelli, 22
47814 Bellaria – Igea Marina (RN)
Italy

natacha.niemants@unimore.it
nati.niemants@gmail.com

About the author

Natacha S.A. Niemants is a free-lance conference interpreter in Italian, French and English, completing a PhD project at the University of Modena and Reggio Emilia on the training of public service / community interpreters. She is participating in an interdisciplinary research team focusing on interlinguistic and intercultural mediation, and her research interests include the following: interpreting in healthcare, interpreter training, conversation analysis, role-play, transcription, and TV interpreting.