Making use of transcription data from qualitative research within a corpus-linguistic paradigm: Issues, experiences, and recommendations

Luke Collins and Andrew Hardie
Lancaster University
{l.collins3,a.hardie}@lancaster.ac.uk

Abstract:

We reflect on the process of re-operationalising transcript data generated in an ethnographic study for the purposes of corpus analysis. We present a corpus of patient-provider interactions in the context of Emergency Departments in hospitals in Australia, to discuss the process through which ethnographic transcripts were manipulated to generate a searchable corpus. We refer to the types of corpus analysis that this conversion enables, facilitated by the rich metadata collected alongside the transcribed audio recordings, augmenting the findings of prior qualitative analyses. Subsequently, we offer guidance for spoken data transcription, intended to 'future proof' such data for subsequent reformatting for corpus linguistic analysis.

1. Introduction

Spoken data collected for ethnographic study or similar qualitative analyses may often be usefully re-operationalised as a corpus *sensu stricto* to allow corpus-linguistic methodologies to be applied (see e.g. Angelelli, 2017; Harrington, 2018; Dayrell et al., 2020). However, the transcription practices used to produce the original data can inadvertently create barriers to such re-operationalisation. In this paper, on the basis of experience with one such collection, we propose guidance for spoken data transcription on how to ensure that transcription practices do not generate such barriers, and to make sure that subsequent use as a corpus will be possible, if desired.

We assume that this re-operationalisation will involve mapping transcripts to some standard machine-readable format, such as XML,[1] in a *structured* manner, i.e. retaining via appropriate markup: textual structure (e.g. utterances); any observer notes, metacommentary, or other contextual information; and text/speaker metadata. We assume further that this mapping will be performed automatically, so the transcription conventions must be unambiguous from the perspective of a conversion program. Rather than using XML (or the like), which is a 'cumbersome' (Love et al., 2017:338) format for direct data entry, we recommend minimal amendments to existing practice to facilitate later automated processing.

These recommendations emerge from our work, described in §3.2, to render utilisable in corpus form a body of transcript data collected for a qualitative, ethnographic study of communication in hospital emergency departments (EDs), some background on which is given in §3.1, after a brief summary of certain key issues in transcription overall (§2). Recommendations for future practice are discussed in §4 and listed in the Appendix.

[1] Other structured formats than XML exist that would be appropriate targets for this type of conversion. One example is the CHAT format used by CHILDES/TalkBank (MacWhinney, 2000). However, any two fully structured formats are trivially interchangeable. To simplify matters, therefore, we will continue to assume an XML target.

2. Contrasting approaches to transcription

Transcription processes are 'variable', significantly influenced by the intended analysis or 'research function' (Bucholtz, 2007). In studies of spoken discourse, there is broadly a distinction between (a) Conversation Analysis (CA) transcripts, which capture phonetic/prosodic aspects of speech, e.g. intonation and allophony (typically via the highly-influential Jeffersonian system: Jefferson, 1983); and (b) orthographic transcripts, which capture little phonetic detail and therefore are suited to investigations into 'morphology, lexis, syntax, pragmatics, etc.' (Atkins et al., 1992:10). Of course, it is possible for phonetic/prosodic detail to be added to orthographic transcripts to support CA, as Rühlemann (2017) has shown with the British National Corpus (BNC) 1994.

Since both spoken corpus construction and transcription for ethnographic analysis are typically orthographic-only, they are thus compatible in principle. Furthermore, qualitative and/or ethnographic transcription data typically includes rich metadata on speakers and communicative contexts, the lack of which has been a criticism of corpus analysis; the inconsistent availability of speaker metadata has been raised as a limitation of the BNC1994 (Lam, 2009) for example. Clearly, then, conversion of transcriptions collected for qualitative research to corpus data should preserve, and make usable, all contextual metadata.

3. A case study: *Communication in Emergency Departments*

3.1. The original data

The collection of transcripts by the *Communication in Emergency Departments* project, led by Diana Slade, ANU Institute for Communication in Health Care, was for the purpose of a study combining 'discourse analysis of authentic interactions between clinicians and patients; and qualitative ethnographic analysis of the social, organisational, and interdisciplinary clinician practices of each department' (Slade et al., 2015: 11). Slade et al. (2015: 1-2) describe this data as consisting of:

> communication between patients and clinicians (doctors, nurses and allied health professionals) in five representative emergency departments in New South Wales and the Australian Capital Territory. The study involved 1093 h of observations, 150 interviews with clinicians and patients, and the audio recording of patient–clinician interactions over the course of 82 patients' emergency department trajectories from triage to disposition.

This 'patient journey' dataset comprises 1,411,238 tokens and supported 'one of the most comprehensive studies internationally on patient-clinician communication in hospitals' (Slade et al., 2015: 2). One major application for such data is in investigations into 'critical incidents' (avoidable patient harm), in which poor health outcomes may be attributable to communication problems. In such cases, typically no record of these problems is available when the health outcomes are investigated, often months after the event.

The transcription of the data was orthographic, showing consideration for topics that any such scheme must address – data protection/anonymisation; the structure of talk (utterance

overlaps); unclear material; non-standardised forms; transcriber notes. These were recorded as follows:

> We have transcribed clinician–patient interactions using standard English spelling. Nonstandard spellings are occasionally used to capture idiosyncratic or dialectal pronunciations (e.g. *gonna*). Fillers and hesitation markers are transcribed as they are spoken, using the standard English variants, e.g. *Ah*, *uh huh*, *hmm*, *mmm*.
>
> What people say is transcribed without any standardisation or editing. Nonstandard usage is not corrected but transcribed as it was said (e.g. *me feet are frozen*).
>
> Most punctuation marks have the same meaning as in standard written English. Those with special meaning are:
> … indicates a trailing off or short hesitation.
>
> == means overlapping or simultaneous talk […]
>
> — indicates a speaker rephrasing or reworking their contribution, often involving repetition […]
>
> [words in square brackets] are contextual information or information suppressed for privacy reasons. Examples:
>
> [Loud voices in close proximity] contextual information
>
> Z1 And your mobile number I've got [number].
>
> (words in parentheses) were unclear but this is the transcriber's best analysis.
>
> ( ) empty parentheses indicate that the transcriber could not hear or guess what was said […]
>
> (Slade et al., 2015:xi)

The transcripts are accompanied by metadata about each participant (role, gender, age, language background, nationality) and headed with information on the context of the 'patient journey' (presenting illness, diagnosis, duration of visit, triage level, number of health professionals seen, and researcher notes).

While Slade et al. (2015:19) present 'information-rich description and analysis', they inevitably explore relatively few instances of interaction in close detail. Given the data's extent, corpus-based techniques can clearly enhance the analysis of (in)effective clinician-patient interaction using this resource. The team working towards this end[2] utilises Lancaster University's CQPweb server (see Hardie, 2012) as its primary analysis platform. As a first step, it was necessary to convert this *ED Corpus* to an XML-based format (for subsequent tagging and indexing) along with *structured* metadata usable within CQPweb (i.e. not free prose metadata). We turn now to the processes involved in this conversion, and problems

[2] This undertaking is a collaboration between the *Emergency Communication* research team, now based primarily at the University of Technology, Sydney Australian National University, Institute for Communication in Health Care (ICH); and the *ESRC Centre for Corpus Approaches to Social Science* (CASS) at Lancaster University.

arising from the original transcription practice (which, obviously, could not have been anticipated).

3.2. Corpus conversion

The transcripts were created, and initially analysed, using Microsoft Word. We used a two-step process to generate usable corpus data. First, the documents were exported to HTML using the relevant Word function (scripted to run automatically across multiple documents using Microsoft's *Visual Basic for Applications*). Plain text export would lose most if not all of the formatting that indicates document structure; HTML export retains this information, albeit not in a form usable by corpus software. Therefore, the second step was to use a bespoke text-conversion script[3] to (a) simplify Word-output HTML in order to (b) parse the document for metadata and utterance content which is then (c) reformatted into XML (per the recommendations of Hardie, 2014). This script also identified unparsable regions of the content, helping us iteratively extend the script to handle them. Our goal was to avoid manual editing while retaining as much information as possible. Achieving this was complicated by two issues: ambiguity and inconsistency.

Ambiguity arises from the use of the same form of notation to represent functionally different content. In the ED Corpus, square brackets were used for 'contextual information or information suppressed for privacy reasons' (Slade et al., 2015). A human reader able to discern intent can easily tell which of three functions (anonymisation, recording non-speech sounds, transcriber comments) applies. A computer cannot. To represent each functional type of information distinctly in the corpus, we had to disambiguate use of square brackets, by automatically compiling a list of instances of square-bracketed text and then coming up with, and coding as regular expression tests, rules to distinguish anonymised speech from transcriber comments (etc.).This process made it apparent that square brackets were also used to code *embedded turns*: short turns not represented as separate utterances but rather as events within the utterance of another speaker. Most utterance breaks in the original documents are indicated by the start of a new table-row. This representation was used largely for instances of *backchannelling* (Yngve, 1970:568) as follows:

> P No. Just this – just this morning, in the – in the night. I know it coming, the phlegm black. **[D Mm mm]** No more.

This information about backchannelling disrupts the organisation of utterance boundaries, which are essential for automatic identification of speech by specific (kinds of) speakers. Merging this representation with normally-encoded utterance breaks was therefore a priority. But this was hindered by the embedded turns' ambiguity with the three square-bracket functions already discussed. Not even the speaker ID *incipit* ('D' above) distinguishes embedded turns, since transcriber comments can also have this form, e.g. *[D examines patient]*. Disambiguating this required both automatic checks (for obvious backchannel content like *mm mm*) and manual effort to list all non-obvious cases.

[3] Any scripting language (Python, Ruby, Perl, etc.) is suitable for this purpose; we used PHP. The script is based on regular expression search-and-replace to translate existing features of the textual format to XML and is, thus, computationally trivial; its writing required detailed knowledge of the data, and it is thus specific to this corpus and not reusable (though of course the general technique *is* applicable, and has been applied, to other datasets).

Inconsistency is present in multiple aspects of transcribers' practice, including use of punctuation, and the descriptive wording within different types of notes. Some variant practice (for instance, use of round vs. square brackets) actually violated the guidelines. Since Slade et al.'s (2015) subsequent qualitative analysis was *not* impeded, such inconsistency evidently poses no problem for humans. However, corpus methods rely 'on the recurrence of consistent representations of linguistic phenomena' (Adolphs and Carter 2013: 155), i.e. representations that *a computer can recognise as identical*. Keeping each transcriber consistent with others' practice, and with their own over time, is not easy. Expanding the detail in the transcription scheme is no answer, since doing so increases the time needed for transcription (and thus the cost), and in fact makes it *harder* to maintain inter- and intra-transcriber consistency (see Love et al., 2017).

With respect to square-bracket-marked anonymisation alone we observed over 120 different ways of representing omitted names, including:

- Simple placeholder: *[name]*
- Postmodification expressing role: *[Name of researcher], [Name of Father], [Name of female nurse]*
- Speaker ID codes: *[name of N2], [D1]*
- Premodification narrowing reference: *[Male name], [Middle Name Last Name]*
- Various representations of spelt name: *[Spells out name], [S-U-R-N-A-M-E]*
- Commentary on editorial or speaker naming act: *[name removed], [gives first name only]*.

Such inconsistency is not *purely* negative; some of the above represents transcribers' creative exploitation of the tool at hand (square brackets) to encode contextual information not otherwise available to the reader, such as roles and relations of the person mentioned (which otherwise are opaque after anonymisation). Yet still it generated obstacles to conversion into automatically manipulable corpus format. Direct links exist between the information extracted from the transcripts and specific query tools in CQPweb, as follows (other software presents intersecting but not identical sets of affordances):

- *Information on speakers* alongside clear (XML-format) *utterance boundaries*: CQPweb can restrict queries to utterances of speakers with particular metadata features (e.g. sex, age, or more relevantly role as patient, doctor, nurse, etc.),[4] as well as generate statistical comparisons (e.g. keywords) between subcorpora delineated by this information.

- *Information on contextual features*: CQPweb can treat each interaction as a 'text', and contextual information as text-level metadata, enabling division of the corpus by these factors (such as timing of different stages of the recorded 'patient journey') as well as, or instead of, speaker features.

- *Anonymisation and vocalisation labels*: if consistent, these can be located using CQPweb's query syntax, allowing, for instance, language use in the vicinity of

[4] Hardie (forthcoming) explains the systems used within CQPweb to accomplish this.

laughter, or contexts where personal information is expressed at high density (and so on), to be identified and studied.

Our experience has been that enriching the corpus with the non-textual information within the original transcription enables analysts to engage with both the 'broad and local sense of context' (Cicourel, 2014: 377) and thus bring corpus and ethnographic approaches closer together. We were thus driven to reflect on how the barriers to this kind of work might be lowered for a wide range of scholars. Problems of transcriber inconsistency and ambiguity are not insuperable, but *are* an impediment. We thus proceed to recommend *slight* modifications to transcriber practice in qualitative/ethnographic research that would make the resulting dataset more easily usable as a corpus down the line. Being based on work with a single dataset, these recommendations may not address facets of the problem at hand that might emerge in other circumstances. Nevertheless the facets they *do* address are sufficiently generic to assure at least some wider applicability.

4. Recommendations

4.1. Orthographic consistency

We see no upside to recommending changes in transcription practices substantial enough to impede the initial purpose of qualitative data transcription. Instead, we suggest tweaks to existing practice to 'future-proof' such datasets so that their subsequent use within a corpus-linguistic paradigm is facilitated: by enhancing consistency, reducing ambiguity, and thereby making data conversion more straightforward and reliable.

On the consistency front, the key is to minimise what Andersen (2016) calls 'unmotivated variability' in transcription (reproduction of actual variation in the language being represented is motivated and of research interest, of course). The items potentially affected are largely those lexicalised forms and semi-lexicalised forms (Andersen, 2016) for which there is some orthographic convention but no strong standard. Colloquial pronunciations often have multiple potential orthographic representations, as do vocalised or filled pauses (*um*, *erm*, *uh*, and friends). Atkins et al. (1992) recommended establishing a closed set of non-standard forms that transcribers are permitted to use, and this remains best practice. Without such limits, it is impossible to know whether word-forms *er*, *uh*, *ehhhn* (etc.) represent phonetically different vocalisations, or whether there is a real distinction between the weak-form of *have* represented as *of* versus *'ve*, without listening to the recording. A pilot stage may be required to understand exactly what kind of variation transcribers need, and are able, to utilise consistently, as demonstrated by recent work on spoken corpus creation (Love et al., 2017; Gablasova et al. 2019). Establishing a closed set reduces (but does not eliminate) the scope for inter- and intra-transcriber inconsistency.

4.2. Unambiguous markup of non-speech material

To limit, and if possible eliminate, ambiguity of notation, we recommend that the conventions should present *clearly distinct* representations for distinct kinds of comment or label, all of which must also be unambiguously distinct from actual spoken content. This is partly accomplished by Slade et al.'s (2015) system (see §3.1), which mandates use of round

brackets to mark spoken content as unclear, in contrast to the square brackets' functions. In practice, this distinguishes:

1. They asked me my (name?)
2. [name] is here to see you.

In (1), the speaker has said *name*, but the speech is unclear (perhaps because of audibility issues); the round brackets and question mark express the transcriber's uncertainty. In (2), the speaker has used some person's actual name; the square-bracketed label records and classifies the redaction. This represents good practice but not *best* practice, as under these guidelines multiple distinct kinds of insertion are delimited by square brackets (vocalisations, transcriber comments, embedded utterances, and redactions for privacy), and this proved a substantial impedance to corpus conversion. Recording functionally distinct information types in unambiguously distinct forms requires only a minor adjustment to transcription practice.

We suggest that different uses of square brackets should be indicated by a *flag* character directly after the opening square bracket, with the same principle applied to round brackets if they have multiple uses. Individual punctuation characters such as number-sign/hash (#), at-sign (@), colon or semi-colon are recommended,[5] since transcribers are unlikely to begin bracketed material with any of these and each flag expresses directly what the square brackets represent, e.g.:

- Vocalisation: **[@laughs]**
- Transcriber comment: **[#D fills in form]**
- Embedded utterance: **[=P yeah]**
- Redaction: **[name]**

The presence of flags in the original transcript will not impede manual analysis, but does make it entirely mechanistic to automatically convert the above to XML or another structured format, for instance:

- **[@laughs]** becomes `<voc desc="laughs"/>`
- **[#D fills in form]** becomes `<comment content="D fills in form"/>`
- **[=P yeah]** becomes `<u who="P" trans="overlap">yeah</u>`
- **[name]** becomes `<anon type="person"/>`

The precise flags used can be adjusted per the requirements of any particular data collection or corpus conversion endeavour. The example XML above was devised in light of our particular needs: (a) to have this non-textual information accessible via CQPweb, into which non-linguistic data can only be input in the form of simple pseudo-XML tags; (b) to exclude from corpus queries and word counts the content of sometimes lengthy non-speech material in transcriber comments. Since CQPweb models text as a sequence of tokens, where XML takes up no space, but sits *between* tokens, we generate a dummy token unmistakeable for any real word, i.e. `<anon>--anonname</anon>` (where *--anonname* need only be unmistakeable for any real word) to allow redacted word(s) to take up space in the token

[5] To avoid complications, it is better not to use as a flag any symbol with a special meaning in *regular expression* syntax, since most query engines interpret these in special ways. This includes question mark, plus, asterisk, circumflex, and dollar-sign.

sequence. The ability to vary the data representation via automatic conversion is a further advantage of the adjusted practices we suggest.

4.3. Standardised values and comments

Searchability and countability of redactions, vocalisations and the like are further enhanced if their descriptions, the material *within* the brackets, are presented in regular form. Frequent in our corpus are vocalisations 'laughter' and 'coughing', features relevant to the research aims (e.g. analysing, respectively humour and rapport-building/illness and audibility). However, the transcribers variously use nouns, plain verbs and third-person verbs to code these: *[laughter]*, *[laugh]* and *[laughs]* all occur. We recommend that transcriber practice should standardise on just one style, e.g. **[@laughs]**, **[@coughs]**. Ideally, a closed list of permissible vocalisation descriptions should be defined.

The same principle applies to redactions. Using the standardised format **[name]**, other types of anonymised information (e.g. dates of birth, telephone numbers) can be given defined labels from a restricted set. For some projects, a single category for all proper nouns might suffice; in other cases, separate labels for **[name]**, **[place]**, **[organisation]**, might be needed. As an extension, ID codes for speakers can be permitted, for mentions of catalogued discourse participants, e.g. **[P]**, **[D2]**. These can be automatically recognised, and used to create an `<anon/>` which records who has been mentioned, e.g. as `<anon type="person" who="D2"/>`. Finally, the same closed-list approach should be applied to categorical values taken by metadata on texts or speakers, so that it may be automatically extracted in structured form.

Using closed lists has the *dis*advantage of denying transcribers freedom to record on their own initiative relevant but non-predefined information. For example, a transcriber might wish to record an anonymisation as **[name of patient's mother]** to aid analysts' understanding of the text, but our recommendations would require just **[name]**. To counteract this, we recommend the non-standardised comment mechanism, that is **[*#text-of-comment*]**, where comments directly after some other element are understood by convention to relate to it: **[name][#patient's mother]**. In the same way, vocalisation descriptors can be enhanced in an adjacent comment, e.g. **[@coughs][#to draw doctor's attention]**, to support subsequent investigations of form or function. We strongly suggest not limiting use of comments in any way, retaining this one notation as a highly flexible space for *any* contextual information the transcriber thinks pertinent.

Finally, we recommend a fairly informal approach to silences. Transcribers should not be tasked with precise measurements of pauses. In the ED transcripts, ellipses were used for either a pause or the trailing-off of an utterance. We recommend that ellipses **…** be restricted to indicating a discernible pause of less than (roughly) three seconds, for use *within* a turn, not at the end. For silences of three to ten seconds, we suggest the convention **[pause]**, and for prolonged silences of more than ten seconds, the convention **[silence]**. These latter two conventions are not ambiguous with redactions because *pause/silence* are not types of anonymisation.

5. Conclusion

The above recommendations are designed to generate minimal ambiguity when qualitative-research transcription data is mapped to XML or other structured format and operationalised as a searchable corpus. The Appendix presents their implementation as modifications to Slade et al.'s (2015) conventions; however, what we really wish to underline are the advantages *in principle* of conventions that are utterly unambiguous, and thus manageable by computer programs.[6] Defining and enforcing such conventions simplifies and regularises transcription, while permitting rich contextual information via the flexible transcriber comment mechanism.

Acknowledgements

This research was supported in part by the Economic and Social Research Council, grant reference ES/R008906/1, *ESRC Centre for Corpus Approaches to Social Science (CASS)*.

Appendix: Summary of recommendations

Non-standard variants, filled pauses, and weak forms should not be transcribed impressionistically. Rather, researchers should pre-define a closed set of allowed lexical and semi-lexical items of these kinds, based on considerations of conventionalised orthography (e.g. *erm, dunno*) and of the range of variation present in the data and pertinent to their research interests.

Punctuation use should be minimal. Question and exclamation marks may be left to the transcriber's intuition regarding whether intonation cues require them (rather than explicit interrogative/exclamative forms).

(*word*?)	Round brackets/parentheses should indicate a transcriber's best guess at unclear words on the recording; an optional question mark can indicate especially uncertain guesses; and empty brackets **( )** a totally uninterpretable word(s).
[name]	Square brackets with no flag should indicate content that has been redacted (with a label such as *name*, *place*, *dateofbirth*, as appropriate)
[#*comment*]	Square brackets flagged with hash should indicate transcriber comments: any observations, interpretations or descriptions that are not actual recorded speech.
[@laughs]	Square brackets flagged with @ should indicate vocalisations (with a label such as *laughs*, *coughs*, *groans* …) that occur within an utterance.
…	Short pause: less than three seconds.
[pause]	Medium pause: 3-10 seconds.
[silence]	Long pause: anything more than 10 seconds.

[6] Unambiguous markup also makes possible automated detection of mistakes made in transcribers' use of brackets and other notation; we lack space to explore this issue in detail, however.

=	Onset of a turn which overlaps the prior turn.

References

Adolphs, S. and Carter, R. 2013. *Spoken Corpus Linguistics – From Monomodal to Multimodal*. London: Routledge.

Andersen, G. 2016. 'Semi-lexical features in corpus transcription: Consistency, comparability, standardisation', *International Journal of Corpus Linguistics* 21(3), pp 323-347.

Angelelli, C. V. 2017. 'Can ethnographic findings become corpus-studies data? A researcher's ethical, practical and scientific dilemmas', *The Interpreter's Newsletter* 22, pp 1-15.

Atkins, A., Clear, J. and Ostler, N. 1992. 'Corpus design criteria', *Literary and Linguistic Computing* 7(1) pp 1–16.

Bucholtz, M. 2007. 'Variation in transcription', *Discourse Studies* 9(6), pp 784-808.

Cicourel, A. V. 2014. 'The interpenetration of communicative concepts: examples from medical encounters', in H. E. Hamilton and W.-Y. S. Chou (eds.) *The Routledge Handbook of Language and Health Communication*, pp 375-388. London: Routledge.

Dayrell, C., Ram-Prasad, C. and Griffith-Dickson, G. 2020. 'Bringing corpus linguistics into Religious Studies: Self-representation amongst various immigrant communities with religious identity', *Journal of Corpora and Discourse Studies* 3, pp 96-121.

Gablasova, D., Brezina, V. and McEnery, T. 2019. 'The Trinity Lancaster Corpus: Development, description and application', *International Journal of Learner Corpus Research* 52, pp 126-158.

Hardie, A. 2012. 'CQPweb – Combining power, flexibility and usability in a corpus analysis tool', *International Journal of Corpus Linguistics* 17(3) pp 380-409.

Hardie, A. 2014. 'Modest XML for Corpora: Not a standard but a suggestion', *ICAME Journal* 38, pp 73-103.

Hardie, A. Forthcoming. 'Managing complex and arbitrary corpus subsections at scale and at speed: from formalism to implementation within CQPweb'.

Harrington, K. 2018. *The Role of Corpus Linguistics in the Ethnography of a Closed Community: Survival Communication*. New York: Routledge.

Jefferson, G. 1983. 'Issues in the Transcription of Naturally-occurring Talk: Caricature vs. Capturing Pronunciational Particulars', *Tilburg Papers in Language and Literature* 34. Available at: http://liso-archives.liso.ucsb.edu/Jefferson/Caricature.pdf.

Lam, P. 2009. 'The making of a BNC customised spoken corpus for comparative purposes', *Corpora*, 4(1), pp 167-188.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. 2017. ‘The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations’, *International Journal of Corpus Linguistics*, 22(3), pp 319-344.

MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum.

Rühlemann, C. 2017. ‘Integrating Corpus-Linguistic and Conversation-Analytic Transcription in XML: The Case of Backchannels and Overlap in Storytelling Interaction’, *Corpus Pragmatics* 1, pp 201–232.

Slade, D., Manidis, M., McGregor, J., Scheere, H., Chandler, E., Stein-Parbury, J., Dunston, R., Herke, M. and Matthiessen, C. M. I. M. 2015. *Communicating in Hospital Emergency Departments.* Heidelberg: Springer.

Yngve, V. 1970. ‘On getting a word in edgewise’, *Papers From the Sixth Regional Meeting of the Chicago Linguistic Society*. pp. 567–77. University of Chicago.