

BUILDING A CORPUS OF SPONTANEOUS INTERACTION⁷

N. J. Enfield, Kobin H. Kendrick, J. P. de Ruiter, Tanya Stivers & Stephen C. Levinson

Project	Interactional Foundations of Language, Categories across Language and Cognition.
Task	Collect high quality video recordings of spontaneous, naturally-occurring interaction for transcription.
Goal of task	To acquire a corpus of video data, for investigating the underlying structure(s) of interaction cross-linguistically and cross-culturally.
Prerequisites	Access to, and familiarity with, informal settings for conversation in a speech community; good command of the language, access to consultants to help with transcription and translation of recordings.
Outcome	Many of the research activities of the Interaction project depend on having an extensive and varied corpus of social interaction in your language. This task is an important prerequisite to the major comparative projects being undertaken in coming years.

Background

Research on video and audio recordings of spontaneous naturally-occurring conversation in English has shown that conversation is a rule-guided, practice-oriented domain that can be investigated for its underlying mechanics or structure (for overviews see Levinson, 1983; Heritage, 1984; Sidnell, 2010). Systematic study could yield something like a grammar for conversation. Over the past 40 years, a variety of practices and structures have been identified, including the following:

- 1) A system for TURN TAKING in conversation (Sacks, Schegloff & Jefferson, 1974) has held up to moderate cross-linguistic investigation.
- 2) A description of SEQUENCE ORGANISATION in conversation has proven critical in identifying alternative organisations of overall structure in conversation (e.g., story telling) and in other speech exchange systems (e.g., news interviews or teacher-student encounters) (Schegloff & Sacks, 1973; Schegloff, 2007).
- 3) An outline of REPAIR in English has identified practices for managing problems of speaking, hearing, and understanding in conversation (Schegloff, Jefferson & Sacks, 1977; Schegloff, 1979; Schegloff, 1992).
- 4) A description of STRUCTURAL PREFERENCE has shown how the design of a turn can facilitate a particular response such as a “yes” (e.g., “Are you going to the party?”) or a “no” (e.g., “You’re not going?”) (Sacks, 1973; Heritage, 1984; Pomerantz, 1984; Raymond, 2003). This has provided insight into how interaction is fundamentally organised, and also has been utilised in applied situations such as when communication fails to work well in a medical consultation.

Much existing research has relied on telephone calls between English-speaking participants. This project seeks to address two primary gaps in current research:

⁷ This revised version supersedes all previous versions (e.g., Field Manual 2010).

- 1) Social interaction primarily occurs, and is arguably designed for, face-to-face contexts where people have visual access to each other's behaviour, and to the common environment. Video recordings of face-to-face interaction provide access to the multimodal aspects of communication which play a role in any ongoing interaction (and hence in any interactional/linguistic practice).
- 2) Principles outlined in social interaction research to date have been claimed to apply universally. For example, the Sacks et al. (1974) model of turn-taking has been assumed, until proven otherwise, to operate in all languages and cultures. When languages other than English have been investigated, they have tended to be other European languages. This project aims to investigate structural properties of language use in typologically, areally and genetically diverse languages.

Research questions

What are the principles that underlie the structure of social interaction? Are these principles the same cross-linguistically/cross-culturally?

Task

The task includes two phases: first, the field researcher records episodes of spontaneous interaction and then, second, works with consultants to transcribe it.

Recording

Field researchers are asked to video-record a range of different “maximally informal speech events” involving a range of different participants. Maximal informality is defined as the situation in which the fewest structural constraints on interaction apply. This generalisation is meant to discriminate a maximally informal “genre” from explicitly restrictive environments such as ceremonies, speeches, trials, interviews, requested story tellings, etc. The equivalent of “hanging out”, “gossiping”, “chatting”, or “doing nothing” would qualify. More informal situations might be identifiable in terms of *activities*: casual conversation has no explicit pre-determined goals, is often embedded in other activities (like peeling potatoes), is what you do while waiting to do something more important, and doesn't need elaborate initiation or termination. One could also use the *participants* as a clue: e.g., the kind of verbal activity characterising same-sex teenagers of the same hamlet in an idle moment. In particular, the following tend to apply in the kinds of situations we are after:

- a) who the participants are is not pre-determined
- b) the dialect or register is not pre-set to be formal
- c) the order of taking turns at talking is not pre-allocated
- d) what can be said is not pre-determined (e.g., unlike in a marriage ceremony)
- e) where and how the participants are spatially positioned is not pre-determined

Note that the most informal situations are not necessarily the most frequently observed. The researcher is not always party to private activities, and may only see more formal ones. Also note (important!), constraints on interaction are never fully absent – we are looking for the *relatively* least constrained speech situation.

These recordings do not have to be very long (5-10 minutes is often enough). Ideally 10 recordings of 10-30 minutes in length involving 2-5 participants per interaction would suffice for an initial corpus from which to draw data for sub-projects.

Conversation is difficult to transcribe and understand, so you need to work with the best quality data possible. It is not easy to collect good quality interactional data. Do not expect that you will succeed in any given attempt to make a recording of interaction. You should expect all sorts of interruptions, technical and other unforeseen problems to thwart your attempts to record, and you may be unable to use many of the recordings you make. If you expect to meet these difficulties then you can aim to record as much material as you can, which will allow you to discard the low quality material, and pick the best material to work with.

A note on visual quality: Please read the instructions for video-recording at the beginning of this manual, and please pay special attention to exposure and to composition of the frame. By ‘exposure’ we mean getting the settings right for the level of lighting available. Try to avoid situations in which speakers are in dark areas where the background is bright. If you must film in such a situation, make sure you set the ‘backlight’ option on the camera. By ‘composition of the frame’, we mean getting certain things in the shot. Do not film close-up shots, as you will miss a lot of important information. People’s whole bodies are important in interaction, especially their hands and arms. You will therefore have to leave enough space in the frame for large/wide gestures not to be cut off. Also, you should try to keep all participants in the shot, even when they are not talking. It is best if you can have the camera set up on a tripod, but if you need to film hand-held, that’s okay too. Just be very careful to keep the camera as steady as humanly possible (e.g., by propping yourself against a post or wall if one is handy). Also, after you have set the frame composition, you should avoid using the ‘zoom’ at all costs. If you are interested in looking at eye gaze, we recommend you use more than one camera, one facing toward each of the participants.

Transcription

Once you have made some recordings, you will need to select sections which are of good quality (i.e., in both audio and visual quality) and work with consultants to transcribe the linguistic material in detail. Rather than transcribe one long recording, you may wish to begin with shorter segments from multiple recordings. Six 10-minute segments from six different conversations will be more useful in IFL subprojects than a single 60-minute segment. For transcription we recommend that you use Elan, if possible. A template with predefined annotation tiers and linguistic types is available on the IFL website.⁸ For convenience, you may wish to do the transcription using the audio signal only – that is, by first exporting the audio from the video file. Note that if you are interested in one phenomenon in particular (e.g., repair), you might search the video data for instances and transcribe just those sequences of talk, including any possibly relevant talk before and after the target item. Most importantly with the transcription is to capture as many details of what is said as possible, including glottal cut offs, stretching of sounds, or other types of hesitations and perturbations, and changes in pronunciation or syntax. When working with consultants be sure that they do not “fix” ungrammatical or odd sounding turns but assist with capturing all details of the talk as produced.

⁸ <http://www.mpi.nl/research/research-projects/interactional-foundations-of-language/tools>

Analysis

The data that you collect and transcribe will be used to investigate interactional practices as linguistic systems from a cross-linguistic and cross-cultural perspective. In the coming years, IFL subprojects will investigate a wide range of domains: other-initiated repair, place reference, requests and recruitments, timing and turn-taking, action formation and ascription, and perception.

Although each subproject will specify its own requirements, many IFL subprojects build on the methods of conversation analysis. A general overview of conversation-analytic methods can be found in Sidnell (2010:20-35) and more in-depth discussions in Schegloff (1996, 1997).

References

- Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press
- Levinson, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press
- Pomerantz, A. (1984). Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes. In J.M. Atkinson, & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. (pp. 57-101). Cambridge: Cambridge University Press.
- Raymond, G. (2003). Grammar and Social Organization: Yes/no Interrogatives and the Structure of Responding. *American Sociological Review*, 68(6), 939-967.
- Sacks, H. (1973). The Preference for Agreement in Natural Conversation. Paper presented at the Summer Institute of Linguistics, Ann Arbor.
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696-735.
- Schegloff, E.A., & Sacks, H. (1973). Opening Up Closings. *Semiotica*, 8, 289-327.
- Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, 53, 361-382.
- Schegloff, E.A. (1979). The Relevance of Repair for Syntax-for-Conversation. In T. Givon (Ed.), *Syntax and Semantics 12: Discourse and Syntax* (pp. 261-288). New York: Academic Press.
- Schegloff, E.A. (1992). Repair after next turn: the last structurally provided for place for the defense of intersubjectivity in conversation. *American Journal of Sociology*, 95(5), 1295-1345.
- Schegloff, E.A. (1996). Confirming Allusions: Toward an Empirical Account of Action. *The American Journal of Sociology*, 102(1), 161-216.
- Schegloff, E.A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes* 23(3), 499-545.
- Schegloff, E.A. (2007). *Sequence Organization in interaction: A Primer for Conversation Analysis*. Cambridge, England: Cambridge University Press.
- Sidnell, J. (2010). *Conversation Analysis: An Introduction*. Chichester, UK: Wiley-Blackwell.