



How to Use Comic-Strip Graphics to Represent Signed Conversation

Kristian Skedsmo

To cite this article: Kristian Skedsmo (2021) How to Use Comic-Strip Graphics to Represent Signed Conversation, *Research on Language and Social Interaction*, 54:3, 241-260, DOI: [10.1080/08351813.2021.1936801](https://doi.org/10.1080/08351813.2021.1936801)

To link to this article: <https://doi.org/10.1080/08351813.2021.1936801>



© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 26 Jul 2021.



Submit your article to this journal [↗](#)



Article views: 954



View related articles [↗](#)



View Crossmark data [↗](#)

How to Use Comic-Strip Graphics to Represent Signed Conversation

Kristian Skedsmo

Department of International Studies and Interpreting, Oslo Metropolitan University, Norway

ABSTRACT

This article explores comic-strip-inspired graphic transcripts as a tool to present conversational video data from informal multiperson conversations in a signed language, specifically Norwegian Sign Language (NTS). The interlocutors' utterances are represented as English translations in speech bubbles rather than glossed or phonetically transcribed NTS, and the article discusses advantages and disadvantages of this unconventional choice. To contextualize this exploration of graphic transcripts, a small-scale analysis of a stretch of interaction is embedded in the article. The extract shows conversational trouble and repair occurring when interlocutors respond to utterances produced while they as recipients were looking elsewhere. The NTS extract is introduced with a short sample of multilinear, Jefferson-inspired glossed transcript and then presented in full as graphic transcript. The article concludes that for presenting nonsensitive data, graphic transcripts have several advantages, such as improved access to visual features, flexible granularity, and enhanced readability. Data are in Norwegian Sign Language with English translations.

The aim of this study is to explore comic-strip-inspired graphic transcripts (Laurier, 2014, 2019; Wallner, 2017a, 2017b, 2018) to present research data from multiperson signed language conversation. An evaluation of the adequacy of a transcript “must be based on specific research goals and particular research questions” (Duranti, 2006, p. 307). A small-scale analysis of gaze direction and conversational trouble in a stretch of Norwegian Sign Language (NTS) multiperson conversation (Bolden, 2011; Egbert, 1997) is therefore embedded in the article. The extract is introduced with a sample of a multilinear Jefferson-inspired transcript but is chiefly presented as graphic transcripts where the utterances are represented as English translations. Each panel (picture) is tagged with the time code of the corresponding frame in the video clip and line numbers referring to the multilinear transcript with glosses. The full multilinear transcript, a printer-friendly version of the graphic transcript, and full-speed and half-speed subtitled video clips are available from the [Open Science Framework \(OSF\)](#) as supplementary material. This way, readers are invited to consider the research question: What are the advantages and disadvantages of graphic transcripts with English translations to (re)present signed language data for conversation analytic publications?

After this introduction follows a short description of the trouble of seeing in signed conversation. The next section will discuss ways of (re)presenting video-recorded conversational data from spoken or signed languages. Special attention is given to situations where data are in a language other than that of the publication and whether utterances in a graphic transcript of signed interaction should be rendered as glosses or translations. The subsequent sections present the data and some methodological and technical issues. Then follows a “test run” with an extract of complex trouble solving in

CONTACT Kristian Skedsmo  kristian.skedsmo@oslomet.no  Department of International Studies and Interpreting, Oslo Metropolitan University, P.O. Box 4, St. Olavs plass, NO-0130, Norway.

I wish to thank Anna-Lena Nilsson, Jessica P. B. Hansen, Jan Svennevig, and three anonymous reviewers for invaluable comments on earlier versions of the text, Beata Slowikowska for L1-consulting, and the participants for generously allowing me to record their private conversations and letting me present them without anonymization.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

multiperson conversation in NTS, presented as a graphic transcript. The extract contains stretches where an interlocutor responds to utterances produced while they were looking away from the addresser. Finally, the use of graphic transcripts will be discussed before the concluding remarks.

Gaze and trouble of perception

In any face-to-face encounter, directions, frequencies, and duration of interlocutors' gaze are considered significant (Kaneko & Mesch, 2013; Kendon, 1967; Kleinke, 1986). In several societies addressers routinely restart their utterances when the addressee is not looking toward them (C. Goodwin & Heritage, 1990). In signed conversation, gaze plays a key role in, e.g., various kinds of reference and verb agreement (Sallandre & Garcia, 2020; Thompson et al., 2006) besides being obviously crucial for the interaction itself. Interactionally, gaze is not only necessary to display interest or to monitor the other's facial expressions and embodied conduct but to perceive what is said. Baker (1977) concludes that addressees in signed interaction must maintain consistent gaze at the addresser and that contributions usually are withheld or repeated until the addressee's gaze is captured. The floor-holder, on the other hand, often does not look at the addressee until the contribution is completed. Then mutual gaze is reestablished to select next speaker and for monitoring feedback. Self-selection is thus only possible when the floor-holder's gaze is directed at the potential self-selector (Van Herreweghe, 2002). These claims have been nuanced as, for example, it has been observed that contributing to the collaborative floor in a multiperson conversation can be given priority over securing that the contribution is seen (Coates & Sutton-Spence, 2001; Kauling, 2012).

In the corpus of informal conversations investigated here, we do not find the strictly organized turn-taking patterns or the efforts to secure common attention that we would expect to find in a job meeting or in a classroom. There are no observable sanctions for toggling visual attention between different schisming conversations (Egbert, 1997), food, drinks, papers, or smartphones. Occasionally, however, such conduct prevents interlocutors from perceiving (parts of) utterances. Other times, overlapping utterances between two interlocutors make an unaddressed participant¹ miss the initial part of the next signer's contribution (Beukeleers et al., 2020).

Although trouble of hearing in spoken interaction and trouble of seeing in signed interaction is comparable in many ways, there are some fundamental differences. Trouble of hearing is often *partial* hearing, i.e., the recipient hears that something is said but not what. Hearing is not as dependent on direction as vision is. If something is expressed outside deaf persons' peripheral vision (Bavelier et al., 2006; Bosworth & Dobkins, 2002; Codina et al., 2011, 2017; Swisher et al., 1989), there is a risk that they will not be aware of it (Johnson, 1991). If the interlocutor does not realize that something is missing, no effort will be made to pursue what was uttered. To build an understanding of what was uttered, these lost parts must be compensated for by comprehension where coherence is constructed from the available pieces (Sanford & Moxey, 1995; Wilkes-Gibb, 1995).

Transcription of face-to-face interaction

To study face-to-face interaction, it is necessary to capture the flow of signals and practices. They need "preserving in some stable form" (Pizzuto et al., 2011, p. 205) for analyses and eventually for presentation to an audience. Many spoken languages have written forms with simpler structures and stricter conventions than what we find for spontaneous face-to-face interaction. The strict and simple "rules" and the static modality make written language more convenient to study than its spoken counterpart. Given that equipment for recording auditory and visual language has only been available for a small part of the approximately 2,500 years of scholarly linguistics, it is understandable that written texts have been its main object (Allwood, 1996; Linell, 1982, 2005).

¹Alternatively *ratified participants* (Goffman, 1963) or *third-parties* (Dyrel, 2014).

Transcription in conversation analysis (CA) attempts to capture the talk “as it is” in its natural habitat (Hepburn & Bolden, 2012; Jefferson, 2004). Phonetic transcription (like IPA), uses specialized symbols, while Jeffersonian and other CA transcription conventions like, e.g., GAT (Selting et al., 2011), use the Latin alphabet along with symbols accessible from a common computer keyboard to capture pronunciation, intonation, pace, volume, voice quality, simultaneous talk, etc. A Jeffersonian transcript of spoken language is relatively readable to readers who know that particular language and will in many cases meet the demand of Pizzuto et al. (2011) by allowing “*anyone who knows the object language* to reconstruct its forms, and its form-meaning correspondences in their contexts, even in the absence of ‘raw data’” (p. 205, original emphasis).

Multilinear transcripts (Hepburn & Bolden, 2012) of spoken language interaction have been developed for investigating and displaying gesture and other visual conduct in research with focus on embodied resources (Heath & Luff, 2012a, 2012b; Heath et al., 2010; Mondada, 2011, 2018, 2019). Multilinearity is also employed to present findings from languages other than that of the publication. One line presents a close transcription of the original language, another consists of a morpheme-by-morpheme representation of words and functions of the original talk translated into the publication’s language, often called “glossing” (Nikander, 2008; Pizzuto et al., 2011; Sallandre & Garcia, 2013). A third line provides a translation into the language of the publication (Hepburn & Bolden, 2012). Multilinear transcripts can be voluminous and difficult to read. As all transcripts, they must balance detail and accuracy against readability.

Transcribing signed languages

Even though there is often a notable divergence between the pronunciation and the standard spelling of words, alphabetically written languages inevitably derive from (some variant of) their spoken counterpart. Signed languages have no established written form (Crasborn, 2014). To transcribe signed languages sign by sign, with an accuracy resembling Jeffersonian transcripts, there are currently two solutions. One is to choose among the different phonetic transcription systems that have been developed since the 1960s, like Stokoe notation, Sutton SignWriting, or HamNoSys (see, e.g., Hoffmann-Dilloway, 2011; Takkinen, 2005). They have been developed by, and are used in, different academic environments (Stone & West, 2012). Phonetic transcripts can provide a high level of detail, conveying precise identifications of the signs and how they are articulated. A challenge so far is the limited number of competent users.

The most common solution in international publications on signed languages has been to present signed utterances as transcripts based on glossing, where each sign is represented with words from a spoken/written language (often English) (Crasborn, 2014; Pizzuto et al., 2011). Glosses are regularly written in upper case in uninflected form (Rosenthal, 2009; Supalla et al., 2017). Grammatical or interactional modifications (plural of nouns, directions of verbs, etc.) are often added with symbols and abbreviations.

Signed languages have many, all-visual, articulators (two hands, mouth, eyebrows, etc.), and discriminating between embodied conduct and “talk” is problematic (Esmail, 2008). Multilinear transcripts (“music-score transcripts,” Manrique, 2016, 2017; Napier, 2007; Van Herreweghe, 2002) are commonly employed, with different articulators presented on different lines.

Spoken language glossing ordinarily constitutes a semitranslated line between the transcription of the original language and the translation to display the function of each morpheme in the first line. Glossing of signed languages is regularly displayed as if it was in itself a sufficient representation of the signed language (Petitta et al., 2013). This tradition is criticized for being assimilationist by emphasizing structural commensurability with spoken languages and hence masking fundamental differences (Pizzuto et al., 2011; Sallandre & Garcia, 2013). Stretches of signing that contain few or no lexical signs but instead make use of nonmanual markers (Valli et al., 2011), classifiers (Emmorey, 2003), and constructed actions (Ferrara & Johnston, 2014) are difficult to gloss in a consistent, brief, and comprehensible way. Another point of criticism is that glossing says nothing about the form of the signs and hence fails to enable readers to reconstruct the original form of the utterances (Pizzuto et al., 2011).

However, transcription is always selective (Hepburn & Bolden, 2012; Hjulstad, 2017; Mondada, 2018; Ochs, 1979), serving the purpose of clearly displaying the specific phenomena of interest for

a model reader (Duranti, 2006; Heath et al., 2010). Duranti (2006) emphasizes that transcripts are “partial and essentialized renditions” (p. 309). Thus, there is no “final”, ‘best’ or ‘only’ way to present the data” (Psathas & Anderson, 1990, p. 78). Another point is to avoid seeing a transcript as the data itself (Psathas & Anderson, 1990). For research on video-recorded conversation, the video files are the data (Hutchby & Wooffitt, 1998). The actual conversation is not available other than in retrospect for those present. Transcription of the interlocutors’ conduct is important for scrutiny during analysis and for presenting the data for readers, but it is not the data, just like René Magritte’s painting of a pipe is not a pipe (Foucault, 1976). Duranti (2006) recites Plato’s allegory of the cave to separate the shadows on the wall (the transcript) from the reality (the data). He comments that while “Plato does not seem to recognize any value in watching shadows, we have made a profession out of it” (p. 306).

When presenting data in a language other than that of the publication, it is useful to consider the benefits of displaying the structures of the original language for the reader. An alternative is to present translations into the article’s language and to show what actions are performed without focusing on how they are expressed in the particular language.

Graphic transcripts

Most adults have experience with reading comics, without convention charts or specific instructions on how to decipher them. A comic strip normally consists of panels representing moments or stretches of time, separated by gutters (Laurier, 2019; McCloud et al., 1994). The drawings can indicate motion by carefully selecting which moments to depict. Drawn arrows, motion lines, or double exposure can also be utilized to add illusions of movement (Eisner, 2001; Laurier, 2014, 2019; McCloud et al., 1994).

Talk in comics is commonly displayed in speech bubbles that, like the panels, are organized left-to-right and top-to-bottom. Different fonts can indicate prosodic features, as can the outline of the bubbles by being, e.g., dotted (whispering) or spiky (shouting) (Eisner, 2001; Kuttner et al., 2020; Laurier, 2019; McCloud et al., 1994; Wallner, 2017a, 2017b). Necessary information not shown clearly in the picture can be displayed in caption boxes in the panels (Kuttner et al., 2020; Laurier, 2014, 2019).

This small selection of comic conventions described here indicates that a graphic transcript can convey information that would require a large number of words to render. However, a major purpose of traditional CA transcription is to display verbal interaction. Photos of signed conversation can convey more information regarding the language production itself than for spoken languages. Still it would take a large number of pictures to capture the complete conversation with a granularity (Mondada, 2018) allowing for precise reconstruction. Conversational data from spoken English can be presented in the speech bubbles as Jeffersonian transcription (McIlvenny, 2014). When presenting interaction in other languages, especially unwritten languages like NTS, to readers assumed not to know it, the speech bubbles must contain transcription/glosses or translations—with differing sets of consequences.

How to (re)present signed language in speech bubbles

As CA traditionally focuses on the surface of talk (Albert & De Ruiter, 2018), representing research on a language by displaying another language, as in this article, might seem like a radical move, or indeed a reactionary one, running the risk of resembling stigmatizing presentations of signed languages as underdeveloped and in need of “naturalization” (Bucholtz, 2000) to be comprehensible to the reader (Rosenthal, 2009). It is hence necessary to emphasize, like Laurier (2014), that these graphic transcripts are designed to present findings (to readers who do not understand NTS) and are less suitable as tools for analytic scrutiny. CA research can have various scientific foci. Grammatical features or pronunciation are not always what the researcher wants to show the reader. Other typical foci can be investigations of communicative actions and practices (Schegloff, 1997), which are generally more translatable than grammatical and pronunciational features.

One of the core qualities of Jeffersonian transcripts is the possibility to render speech and other vocal conduct with sound-by-sound accuracy, including prolonged sounds, false starts, overlaps, etc.

Achieving equal accuracy in a comic-strip-based graphic transcript of a stretch of signed conversation is possible but requires a large number of panels conveying (less than) one sign each, like the drawn representations of Marvel's deaf Avenger character Hawkeye signing (Gustines, 2014). Such fine granularity can be utilized for certain sequences to enhance temporal accuracy and provide the reader with an opportunity for close scrutiny.

The most reader-friendly choice is to present the signed utterances as translations in the speech bubbles. This radically differs from the transcription traditions of CA and deprives readers of the opportunity to know what the interlocutors actually sign in the original language. Figure 1 shows two versions of the first panel of the graphic transcript in this article. The speech bubbles in the left version contain glosses, and those in the right one have translations.

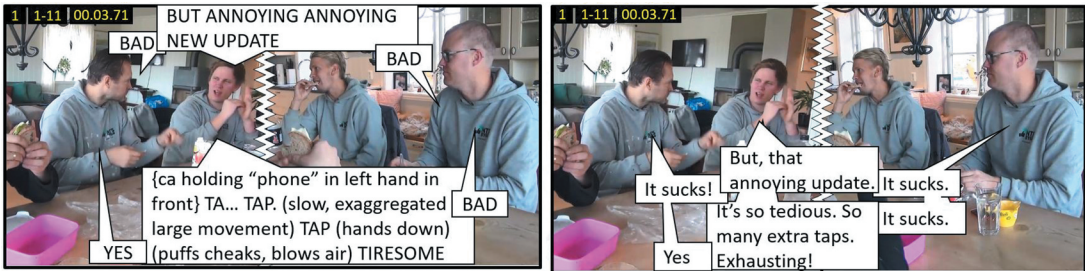


Figure 1. First panel of the graphic transcript with glossing. (See transcription conventions, available from the OSF).

As demonstrated in Figure 1, glossing is possible but reduces readability, and some speech bubbles require more space. For linguistic studies of structural matters, glossing or indeed phonetic transcription would probably be appropriate. For the research focus in this article I have, however, chosen English translations in the speech bubbles.

Depending on the temporal granularity (i.e., the number of panels per second of video), a translated comic-strip format can render more information about the signing itself (and other visual behavior) than pure orthographic glossing normally will. Though not being able to read the original utterances along a timeline is an obvious loss, for certain CA foci it is possible to provide the reader with valuable insights into the actions and practices through translations. However, a crucial point is to consider the *alternatives*. We must remember that glosses are translations too, unidiomatically presented in a sign-by-sign order with symbols and abbreviations added. Another circumstance is that there is always a risk that glossing contributes to a continued prejudiced view of deaf people's languages as "poor" (Rosenthal, 2009; Stone & West, 2012) by presenting their language in a form resembling the ways indigenous people were ridiculed in old comics (Sheyahshe, 2013).

Data and method

The data for this study are extracted from a corpus of a total of 3 h 38 min of informal, multiperson conversation in NTS. The participants are deaf colleagues, video recorded in groups of three to six while having a break at their workplace. Two fixed cameras were set up and left unattended in the room. No tasks were given. The range of topics is wide, and the interlocutors eat, drink, and use their smartphones during the conversations. Informed consent allows me to use transcripts, stills, and video clips from the recordings without anonymizing. All names are pseudonyms.

Even though the recordings constitute the data, they are not as "raw" as video data are sometimes considered. Video and photos are not unmediated, as they are shot from specific angles and often arranged and chosen for specific purposes (Rosenthal, 2009). Still, making excerpts of such data available to the reader can help in approximating the ideal situation "where the reader has as much information as the author, and can reproduce the analysis" (Sacks et al., 1995, p. 27). Achieving

epistemic equality is of course dependent on the reader's knowledge of the particular language. When publishing research on a minority language like NTS, the number of potential readers who understand the language is naturally limited.

Ethical considerations

An obvious set of challenges when publishing photos and video files showing the participants' conduct is related to research ethics. Video-based spoken language ethnomethodology frequently uses anonymized pictures and video files, where faces and voices are made less recognizable (see, e.g., Marstrand & Svennevig, 2018; Mondada, 2019; Wallner, 2017a, 2017b, 2018; Willemsen et al., 2020 for examples). Anonymizing photos or videos of signed language interaction can severely decrease the possibility to discriminate crucial facial actions, mouthings, gaze directions, etc. (Crasborn, 2010); otherwise the anonymizing will appear as symbolic rather than effective (as in Coates & Sutton-Spence, 2001). The participants' generous consent is crucial but not sufficient. The NTS society is a small, vulnerable environment where "everyone knows everyone." The researcher will thus have to balance the value of a clear example against the cost of exposing what might later be experienced as an embarrassing revelation of incompetence, rudeness, etc., by the participants themselves, their friends, or their family.

Graphic transcripts based on frame-grabs from video recordings are nearly as revealing as the videos themselves for recognition of the participants. Still the format allows for ethical considerations. While video extracts might have to be discarded because the participants discuss other people whose privacy the researcher wants to preserve, a graphic transcript allows the transcriber to change names and other references in the speech bubbles to pseudonyms. It is also possible to choose which frame should represent the time sequence of a panel with consideration for how participants appear in each frame.

The graphic transcript in this article

A graphic transcript can be designed in numerous ways to present specific features of the data, as the only transcript or as a complementary transcript together with other transcript formats and/or video. The graphic transcript in this article draws on the comic-strip format, but a number of choices have been made. Any graphic design software (like, e.g., Comic Life or Pixton) not standard on any Windows computer with an MS Office 365 pack was avoided to see what could be done without purchasing special tools.² This choice was made because these applications are not free and hence not accessible to everyone. The panels are frame-grabs from ELAN (Crasborn & Sloetjes, 2008) pasted into PowerPoint for aligning, outlining, cropping, and for inserting speech bubbles, caption boxes, etc. Square speech bubbles are used to save space, and Calibri fonts were chosen over, e.g., Comic Sans MS to reduce associations to humorous comics and the risk of making the interlocutors resemble funny characters. When pictures from simultaneously occurring scenes are combined into one panel, they are separated by a white zigzag line, resembling the diagonal lightning traditionally separating two comic characters having a telephone conversation.³ The completed comic strips were saved in JPG format.

Due to the focus on gaze, one panel does not cover more than one gaze direction of the interlocutor in focus.⁴ (Occasionally several panels cover a stretch of time where gaze is held in one direction.) Speech bubbles are temporally organized vertically over horizontally. This means that the upper bubble precedes the lower. Partly overlapping bubbles indicate partly overlapping utterances, as in Figure 1.

The horizontal widths of the panels do not indicate duration (whereas Eisner 2001 suggests they can, like Laurier, 2019; McCloud et al., 1994). The panel widths in these transcripts are kept to a minimum without hiding important information or giving the impression that there are fewer interlocutors involved than there

²For creating the video extract with two camera angles side by side, a free version of Screencast-O-Matic was used. The free version leaves a logo that is visible on the video and on many of the frame grabs in the graphic transcripts.

³For an example, see <https://no.pinterest.com/pin/32798878533445159/>.

⁴Except Panel 1 (see graphic transcript or Figure 1).

are. Instead, the amount of words in the speech bubbles of one panel gives a hint of the time span of the panel (McCloud et al., 1994). The choice of frame grab to represent the time stretch of a panel is done with regard to how clearly it illustrates the actions conducted. It can be from anywhere in the stretch. Thus, comparing the time codes of the grab does not provide an accurate account of the progression. If the speech bubbles contained glossing, a hash (#) could be inserted, showing exactly what was uttered at the moment the frame represents, as in the conventions developed by Mondada (2018, 2019). Because of the altered syntax in the translations, such markings would be misleading, and overlaps are only approximately indicated. Caption boxes are placed at the top or bottom of the panels. Dotted arrows highlight significant gaze directions, and crucial movements are illustrated with curved motion lines. Figure 2 presents the various features of a panel from the graphic transcript in this article.

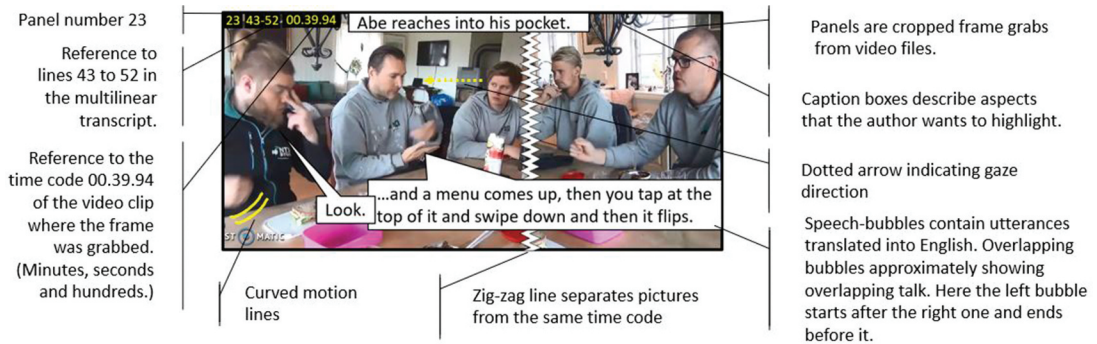


Figure 2. Example of panel from this article with explanations (with curved motion lines and dotted arrow added for demonstration).

Test run with an extract (re)presented as graphic transcripts

This section presents an extract from an NTS multiperson conversation where trouble arises when an interlocutor responds to utterances partly produced while he was gazing away from the signer. First comes a short excerpt of the multilinear, glossed transcript. Then the whole extract is presented as a graphic transcript, piece by piece along the brief analysis.⁵

About the extract

Figure 3 shows the six carpenters Abe, Ben, Carl, Dean, Ed, and Finn having a lunch break. (Dean is not visible in the bulk of the extract because they moved the chairs around after the cameras were set up.) Carl has previously claimed that an iOS update made changing between front and back cameras in FaceTime slow and tiresome and reintroduces this topic at the start of the extract. Ed is seated opposite Carl, while Ben is seated on Carl's right hand side.



Figure 3. Overview of interlocutors from Panel 22 in the graphic transcript.

⁵See supplementary material available from the OSF

Multilinear transcript

The following extract shows the first 13 out of 89 lines of multilinear transcript (available from the OSF). The upper line of the multilinear transcript shows gaze direction. The middle line is glossed signs, and the last line shows English translation. Lines on a common gray background are simultaneous (lines 1–6, 7–8, and 10–11). Hence all actions shown vertically aligned in a gray box overlap.

1.	Carl	Gaze: Ed/Finn-----Ben-----down-front-
		Sign: (smiling) BUT ANNOYING ANNOYING NEW UPDATE
		Trns: But, that annoying update
2.	Ben	Gaze: food-----Carl-----
		Sign: BAD
		Trns: It sucks
3.	Finn	Gaze: food-----Carl-----
		Sign: BAD
		Trns: It sucks.
4.	Ed	Gaze: Abe--food---Carl-----
		Sign: (takes a bite of food)
5.	Abe	Gaze: food-----Carl-----
		Sign: (smiling)
6.	Dean	Gaze: (We cannot see him properly in most of extract)
7.	Carl	Gaze: down front-----
		Sign: {ca holding "phone" in left hand in front} TA... TAP
		Trns: It's so tedious.
8.	Finn	Gaze: Carl-----
		Sign: BAD
		Trns: It sucks.
9.	Carl	Gaze: up front-----down-
		Sign: (slow, exagggregated large movement) TAP (hands down)
		Trns: So many extra taps.
10.	Carl	Gaze: -----Ed-Finn-Abe----
		Sign: (puffs cheeks, blows air) TIRESOME
		Trns: Exhausting!
11.	Ben	Gaze: Carl-----food-Carl-----
		Sign: YES
		Trns: Yes
12.	Abe	Gaze: Carl-----food-----
		Sign: I HAVE-NOT UPDATE HAVE-NOT I WAIT WAIT
		Trns: I haven't updated yet. I'll wait.
13.	Carl	Gaze: Abe-----
		Sign: HEY SOON (become) SLOW HAVE-TO DO-RELUCTANTLY
		Trns: Soon it'll become slow. Then you'll have to.

Graphic transcript with brief analysis


1 | 1-11 | 00.03.71
It sucks!
But, that annoying update.
It's so tedious. So many extra taps. Exhausting!
It sucks.
It sucks.
Yes

2 | 12 | 00.11.09
I haven't updated yet. I'll wait.

3 | 13 | 00.15.46
Soon it'll become slow. Then you'll have to.

4 | 14-15 | 00.18.53
Yeah, when it slows down. It's fast now. I'm happy with it.
Yes, that's right

5 | 16 | 00.20.55

6 | 17-19 | 00.21.71
You tap, and it stays there... forever... then it...

7 | 17-22 | 00.22.85
...flips. Finally.

8 | 17-22 | 00.24.04
Swipe down
Yes.

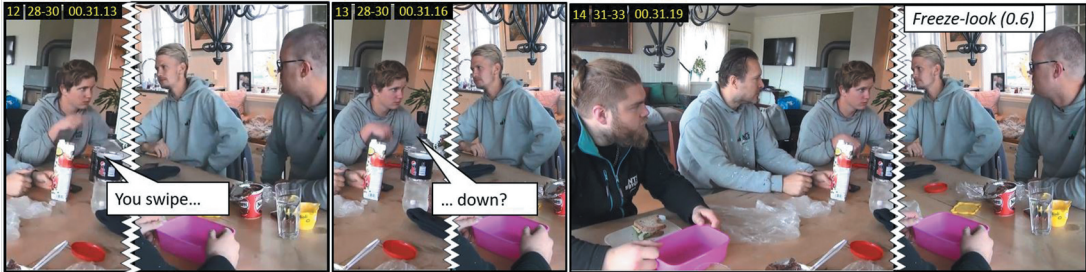
9 | 23-24 | 00.24.78
You tap, swipe down, a menu appears up to the right. You tap there, swipe down and it flips.

10 | 25-27 | 00.28.13
I mean...?
So tedious.

11 | 25-27 | 00.28.83
Flipping slowly back and forth

Panels 1–4 show Carl initiating the discussion about the iOS update. While Carl (5) looks down, Ed and Ben establish mutual gaze, and Ed (6) comments on how slow the process of switching camera is: “You tap, and it stays there . . . forever . . .” Carl (6) looks down during this first part of Ed’s utterance and only looks up toward Ed when Ed (7) says “flips . . . finally” and then (8) “Swipe down.” Simultaneously, Ben (8) says “Yes.” He puts down his carton of milk right next to Carl, and Carl

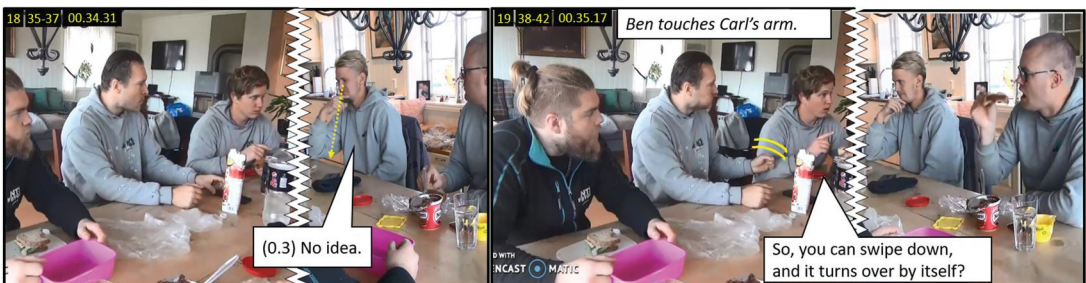
turns from Ed to Ben. Ben (9) describes the process of switching cameras and concludes with a resigned palm-up (10) (Kensy et al., 2018; McKee, 2011). Ed (10) overlappingly comments that it's tedious, but Carl still looks toward Ben. When Carl (11) turns toward Ed, he catches the last part of Ed's utterance, which refers to the display flipping slowly back and forth.

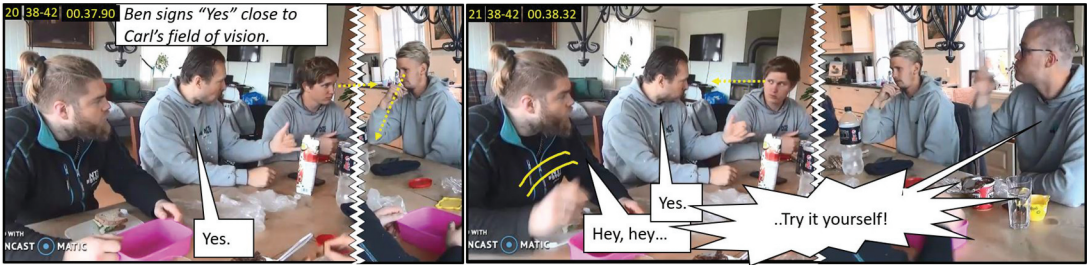


Carl's question to Ed (12–13) indicates that Carl has understood these pieces of talk as (parts of) an explanation of how to switch cameras by merely swiping down. That is a reasonable understanding considering the missed parts from Panel 6 and 10. Ed's initial response (14) to Carl's question (12–13) is a freeze-look response (Manrique, 2016, 2017; Manrique & Enfield, 2015; Skedsmo, 2020a, 2020b). Ed keeps his face and the rest of his body in a steady freeze pose and maintains mutual gaze with Carl for 0.6 seconds before looking down (15).

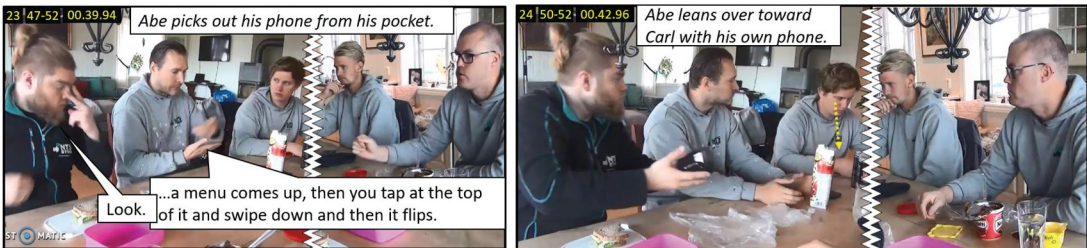
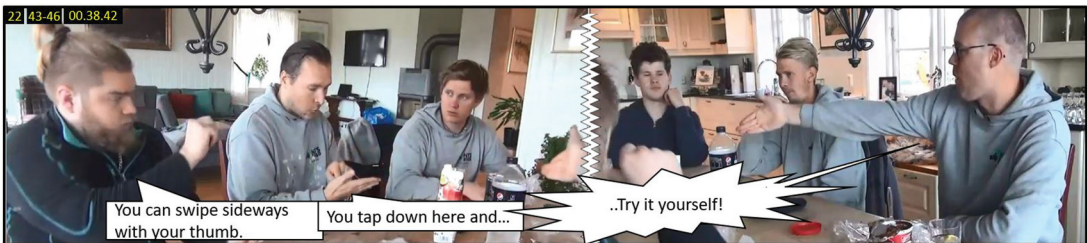


However, Carl is performing a turn-final hold (TFH) (Groeber & Pochon-Berger, 2014), pointing toward Ed (14–15). Ed provides a hedging reply (17)—“You tap and you swipe. I don't know.”—and then looks down adding “No idea” (18).



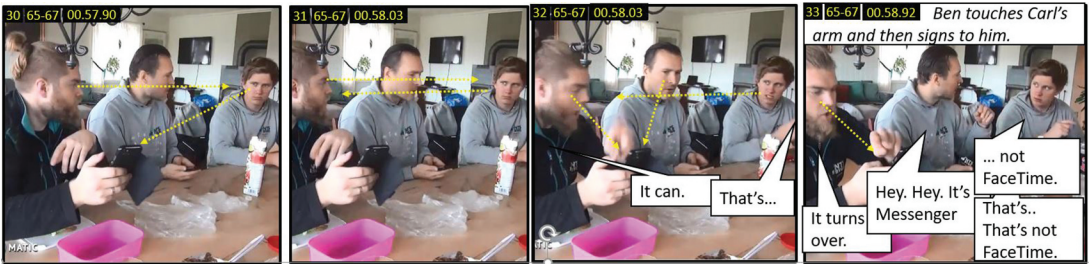
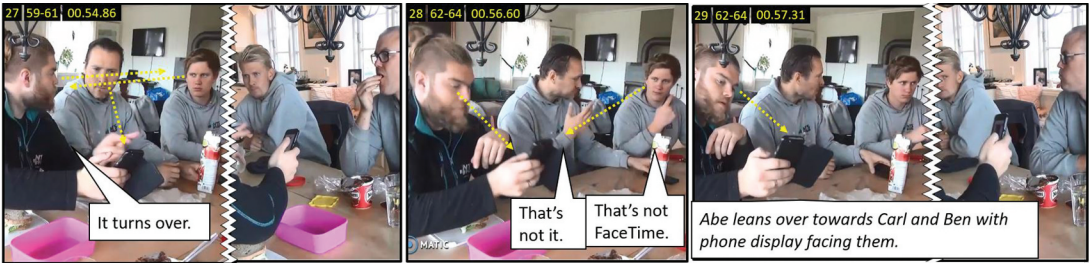


Carl still acts as if Ed has referred to swiping down as a shortcut for switching cameras. When Ed looks back toward him, Carl (19) repeats and elaborates his question: "So, you can swipe down, and it turns around by itself?" Ed again withdraws from the mutual gaze with Carl. Ed's lack of response (20) is instantly followed by three others self-selecting (Lerner, 2003; Sacks et al., 1974). Ben (20–21) confirms and starts an explanation (22). Abe (21) starts summoning Carl, and Finn, with large movements, suggests twice that Carl should try it himself (21, 22).



While Ben (22) starts his explanation, Abe suggests "You can swipe sideways with your thumb," adds "Look" (23), and leans over to get his own phone out of his pocket (24). After Ben has instructed Carl, Abe (25) summons Carl and leans over to show how to swipe sideways.





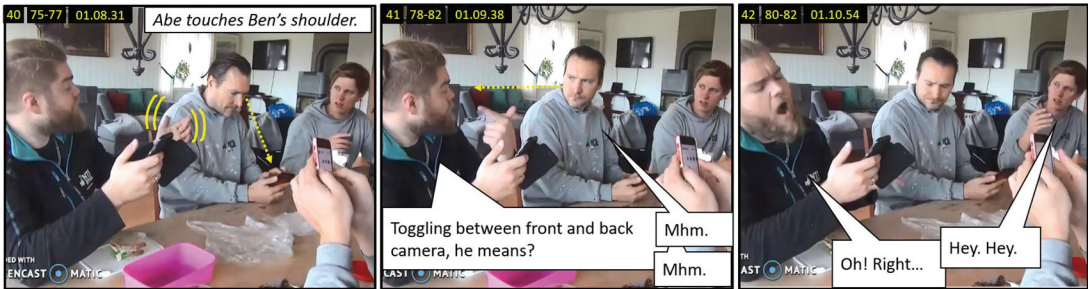
Abe looks toward his own phone during most of the demonstration (25–26, 28–29, 32–33), only with brief glances toward Carl (27, 30). Meanwhile, Carl and Ben (28) say that it is not FaceTime. When Carl looks back toward Abe’s face (31), Abe immediately shifts his gaze back to his phone and continues the demonstration, while Carl and Ben (32–33) comment that what Abe is showing is not FaceTime but Messenger.



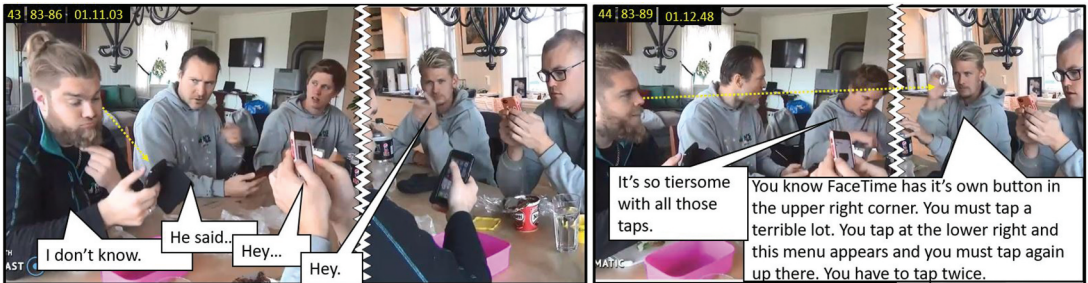
In Panel 34 Abe looks toward Carl, who says, “It’s not FaceTime,” but Abe seemingly takes no notice of it and continues to demonstrate the sideways swiping (35–36).



Ben then (36–38) summons Abe and states that Carl was “asking how to turn the camera in FaceTime.”⁶ Before completing his utterance to Abe, Ben (38) withdraws from their mutual gaze and looks toward his own phone. Abe’s response is a suspension for 0.7 seconds (38) and then a (delayed) change-of-state token (Heritage, 1984), or a display of now-understanding (Koivisto, 2015). Abe (39) shuts his eyes, leans backwards, and says “Oh! Yes. I see!” Ben does not look toward Abe’s display of now-understanding (39). Carl (39) touches Ben’s arm while gazing toward Abe, possibly to guide Ben’s attention toward Abe, but Ben does not respond.



Ben’s newly established broker position (Greer, 2015) overshadows Carl’s ownership of the initial question. Abe seeks to display now-understanding to Ben, not to Carl. Abe (40) touches Ben’s arm, establishes mutual gaze with him, and (41) asks Ben if Carl meant how to toggle between front and back camera. Ben confirms this with two nods (41), and Abe (42) displays now-understanding again in a similar way to what he did when Ben was not looking. Abe’s gaze is now (42) directed toward his own phone, and he probably does not see that Carl summons him.



Abe (43) states that he does not know that. Like the previous time someone failed to answer a question (in Panels 17–22), several others self-select. Carl, Ben, and Ed summon Abe (43). Abe meets Ed’s gaze, and Ed explains the procedure of switching cameras in FaceTime.

We see from the extract that the NTS interlocutors in multiperson conversations cannot and do not always look toward everything that is signed. Looking down makes Carl miss (parts of) utterances, and overlapping turns (10) make it impossible to see both the last part of one utterance and the first part of the next, especially as the two consecutive signers are located so that Carl has to turn his head to look from Ben to the Ed. The evidence for reduced perception lies both in Carl looking down (6) and toward Ben (10) during Ed’s mentioning how tedious it was (except the implicit reference to slowness by Ed [7] saying “finally”) and that Carl asks Ed twice if swiping down will switch cameras.

Partial perceptions of utterances can lead to initiations of repair but do not always (Schegloff et al., 1977). If the perceived parts make sufficient sense on their own, the recipient might not suspect that

⁶This rather mundane act of specifying another’s utterance following an inadequate response is an analytically quite complex practice from a repair perspective. Ben here produces a third-position (Ekberg, 2012; Kitzinger, 2012; Schegloff, 1992), third-person repair (Greer, 2015).

any parts are missing and hence not initiate repair (Skedsmo, 2020b). Both Carl and Abe miss (parts of) utterances as unaddressed participants and not as primary addressees. Being unaddressed might raise the threshold for initiating repair.

There is reason to believe that this kind of fragmented or partial perception of utterances in multiperson conversations is quite familiar to deaf signers and that they are both skilled in, and accustomed to, synthesizing inferential interpretations (Lewandowska-Tomaszczyk, 2017), constructing coherence based on the perceived parts. When motion is detected in peripheral vision, the awareness of it is also likely to contribute to the interpretation. There is a chance that Carl, looking down in Panel 6, has some peripheral perception of Ed's signing. It is even possible that he experiences that he adequately understands what is uttered. Even though peripheral vision extends further horizontally than vertically (Hitzel, 2015), the distance between Ed and Ben⁷ seems to make Carl turn his gaze more than 90 degrees to look toward Ben. However, it is impossible to exclude the possibility of Carl being aware of Ed's signing in Panel 10.

Discussion of the adequacy of graphic transcripts for CA purposes

The analysis and the graphic transcript in this article demonstrate trouble of seeing and examples of NTS signers responding to utterances (partly) produced while they were not looking toward the signer. The data for assessing gaze directions are the video and photos made available to the readers. The number of photos is limited, and two-dimensional photos are not ideal for this purpose. Despite this, photos have been used for determining gaze direction in several studies (Ince & Kim, 2011; Kaneko & Mesch, 2013; Todorović, 2006; Wilson et al., 2000). Multilinear transcripts merely convey the researcher's interpretations of gaze directions. Graphic transcripts, together with the video clips, share the photographic evidence with the reader and therefore make it possible to assess the researcher's analysis.

There is little previous research on deaf signers' gaze directions in actual conversation (Beukeleers et al., 2020). Employing eye-tracking devices in similar informal conversations could reveal numerous details on this matter and help us understand more about signing interlocutors' gaze patterns and how they monitor and relate to the other interlocutors' gaze. Current eye-tracking glasses are very discreet and unlikely to distract signing interlocutors (Beukeleers et al., 2020).

Among the obvious advantages of graphic transcripts is that pictures convey information that would have to be described with numerous lines and words in a multilinear transcription or be dismissed as irrelevant.⁸ For spoken language data, pictures can show context, embodied conduct, and facial expressions that can indicate prosodic features. For signed language data, pictures can capture moments of actual language production.

Laurier (2019) notes that his graphic transcripts lack the temporal precision of a Jeffersonian transcript, but meticulous notation of time can be done where that is in focus. Jeffersonian transcripts render words, sounds, gestures, and other conduct with letters and symbols. They are regularly written with fixed-width fonts like Courier New, and each word or symbol takes up the space it needs independently of duration. Prosodic markings also lengthen the transcript. A stretch of talk marked as produced faster than normal, i.e., a > rush-through<, makes the line longer than if it was produced uttered in normal tempo. There is hence an arbitrary relationship between the length of a line and the stretch of time that it covers (Ten Have, 2007). CA transcripts rather pin down co-occurrences (overlaps, etc.) with a high degree of precision and hence present timing and duration of events relative to each other. Like in traditional comics (McCloud et al., 1994), and indeed in any kind of storytelling, the temporal progression and granularity can vary across the graphic transcript according

⁷The composition of pictures from two camera angles is deceptive, as it looks as if Carl and Ed are sitting next to each other, while they are actually sitting opposite each other.

⁸An anonymous reviewer suggests to "white out" background elements, things at the table, etc., to enhance the focus on what the author wants to show the reader. That is a valid point and would probably work great for reducing visually distracting elements and help the reader focus on what the author wants to show. However, it weakens the argument that graphic transcripts convey a rich impression of the situational context that would otherwise have to be described or discarded as irrelevant.

to what the author wants to highlight. This flexible granularity makes the number of transcription lines corresponding to each panel inconsistent. In Panel 1, the graphic transcript covers 11 lines of transcription and 9.3 seconds of talk, while the four panels 12 to 15 cover only six lines of transcription and as little as 1.6 seconds of video. The pictures in such slowed-down passages can instantly inform the reader about visual co-occurrences, which is especially convenient for visual modes of communication such as gestures or signed language communication. The positioning of the speech bubbles reflects the order of utterances and overlaps. Duration of actions or notable absences of actions can be shown in caption boxes (see Panels 14, 18, and 38).

Not all conversational data are suitable for graphic transcripts. With sensitive data, photos of the participants must be anonymized, e.g., by covering, blurring, or pixelating participants' faces. Especially when working with signed languages, such manipulation risks concealing crucial details of both gaze directions and nonmanual markers and hence reduces the value of the pictures. Often participants can still be recognized by those who know them. Members of the NTS environment typically report that they instantly identify anonymized participants based on very few visual or textual cues (sometimes, of course, mistakenly). For sensitive data, drawn representations, photos of reenactments with other people, or indeed traditional anonymized CA transcripts are possible solutions. However, compared to video clips, graphic transcripts are easier to anonymize with regard to names or other referents mentioned during the conversation that need to be anonymized for privacy concerns. Of the 25–180 frames per second in a digital video, pictures not showing those particular signs can be chosen. [Table 1](#) summarizes advantages and disadvantages of this graphic transcript versus multilinear glossed transcripts.

Publishing CA research with presentations of data in a comic-strip format risks derision as the format is traditionally not associated with science. However, a growing body of scientific publications proves the advantages of various graphic transcripts (e.g., M. H. Goodwin & Goodwin, 2012; McIlvenny, 2014) such as the comic-strip format (e.g., Haddington & Rauniomaa, 2014; Ivarsson, 2010; Laurier, 2013; Wallner, 2017a, 2017b, 2018). However, the most controversial aspect of the graphic transcripts presented here is that the NTS conversation is (re)presented with translations into English. Graphic transcripts can also present structural findings with phonetic transcription or with glossing. For such purposes the temporal granularity will have to be increased to avoid large, cluttered speech bubbles. Among the advantages of translated utterances, readability is the most obvious. Another benefit is that the participants' utterances are displayed in relatively idiomatic language, avoiding the connotations to "broken language" that can be the result of signed language being presented as glosses. This is especially relevant if the reader belongs to the large group of people who do not know signed languages and believe that they are (or indeed that *it is*) less developed than spoken languages. However, there is a risk that members of the NTS community might see the graphic transcript with its reduced renditions of signs and full translations as symptoms of disrespect or oppression and would prefer phonological transcription, glossing, or indeed video. [Table 2](#) summarizes the advantages and disadvantages of choosing English translations over glossed NTS.

Table 1. Advantages and disadvantages of this graphic transcript vs. multilinear, glossed transcripts.

Advantages	Disadvantages
Improved readability for readers outside CA	Unconventional format for CA readers
Improved access to visual features, such as	Less consistent accuracy of the display of relative timing and duration
<ul style="list-style-type: none"> ● gaze directions ● visual, situational context (seating, room) ● communicative actions (signs, summons) ● facial expressions 	Difficult to anonymize without concealing crucial details
Flexible granularity	

Table 2. Advantages and disadvantages of (re)presenting the NTS utterances as English translations vs. glossed NTS.

Advantages	Disadvantages
Improved readability for readers not knowing NTS	Does not show original structure
Less spacious than glossing	Risks oppressing original language.
Does not resemble "broken language"	
Shows translatable actions and practices in a comprehensible/recognizable way	

Concluding remarks

A crucial question when choosing how to present conversational data is what you want to show to whom (Duranti, 2006; Heath et al., 2010; Stone & West, 2012). For presenting findings on gaze directions in NTS multiperson conversation to an audience predominantly consisting of people with an academic interest in conversation but with no knowledge of NTS, graphic transcripts have several advantages. Along with the flexible granularity, the readability, allowing readers without particular experience with multilinear CA transcripts to follow the trajectories, is among the chief gains. Another advantage is increased access to the visual information. The photos do, to a large degree, eradicate the need for describing the physical context, seating arrangements, signs, gestures, facial expressions, etc. These advantages are difficult to accomplish with any effective anonymization, and the graphic transcript is as such only suited for insensitive data where the interlocutors have consented to publication of photos retrieved from the video recordings. Graphic transcripts, as any transcript, allow the use of pseudonyms for people participants mention during their conversations. Choosing among the numerous photos in a video sequence also allows the transcriber to actively avoid less-flattering pictures. The graphic transcripts in this article display translated text instead of transcription or glossing. Other options are available for research questions more concerned with structural or grammatical features, but for investigating particular actions and practices, the readability of the translations can outweigh the gains of glossing or phonetic transcriptions.

The graphic transcripts in this article were created using simple tools available in Office 365. With time I expect that more sophisticated and flexible multimedia interfaces could be developed, e.g., panels showing video clips of the stretch of conversation it covers and flexible speech bubbles that can render different kinds of transcripts or translations or be removed by choice.

Comic-strip-inspired graphic transcripts with translated text seem adequate for presenting findings from various areas of research on face-to-face interaction where the scientific foci are not grammar or other structural issues but instead communicative actions and practices. Offered as the only transcript or as a complementary transcript, their readability may contribute to recruiting new members and future contributors to the fields of CA, interaction analysis, and other research fields employing transcription.⁹

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Oslo Metropolitan University provided funding for this study.

References

- Albert, S., & De Ruiter, J. P. (2018). Repair: The interface between interaction and cognition. *Topics in Cognitive Science*, 10(2), 279–313. <https://doi.org/10.1111/tops.12339>
- Allwood, J. (1996). Några perspektiv på talspråksforskning [Some perspectives on spoken language research]. In M. Thelander (Ed.), *Samspel & variation: Språkliga studier tillägnade Bengt Nordberg På 60-årsdagen* [Interaction & variation: Linguistic studies dedicated to Bengt Nordberg for his 60th anniversary] (23 pp.). Uppsala University, Dept Of Nordic Languages. <http://sskkii.gu.se/jens/publications/docs076-100/081.pdf>
- Baker, C. (1977). Regulators and turn-taking in American sign language discourse. In L. A. Friedman (Ed.), *On the other hand: New perspectives on American sign language* (pp. 215–236). Academic Press.
- Bavelier, D., Dye, M. W. G., & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, 10(11), 512–518. <https://doi.org/10.1016/j.tics.2006.09.006>

⁹See supplementary material available from the OSF

- Beukeleers, I., Brône, G., & Vermeerbergen, M. (2020). Unaddressed participants' gaze behavior in Flemish sign language interactions: Planning gaze shifts after recognizing an upcoming (possible) turn completion. *Journal of Pragmatics*, 162, 62–83. <https://doi.org/10.1016/j.pragma.2020.04.001>
- Bolden, G. B. (2011). On the organization of repair in multiperson conversation: The case of “other”-selection in other-initiated repair sequences. *Research on Language and Social Interaction*, 44(3), 237–262. <https://doi.org/10.1080/08351813.2011.591835>
- Bosworth, R. G., & Dobkins, K. R. (2002). The effects of spatial attention on motion processing in deaf signers, hearing signers, and hearing nonsigners. *Brain and Cognition*, 49(1), 152–169. <https://doi.org/10.1006/brcg.2001.1497>
- Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, 32(10), 1439–1465. [https://doi.org/10.1016/S0378-2166\(99\)00094-6](https://doi.org/10.1016/S0378-2166(99)00094-6)
- Coates, J., & Sutton-Spence, R. (2001). Turn-taking patterns in deaf conversation. *Journal of Sociolinguistics*, 5(4), 507–529. <https://doi.org/10.1111/1467-9481.00162>
- Codina, C. J., Buckley, D., Port, M., & Pascalis, O. (2011). Deaf and hearing children: A comparison of peripheral vision development. *Developmental Science*, 14(4), 725–737. <https://doi.org/10.1111/1467-7687.2010.01017.x>
- Codina, C. J., Pascalis, O., Baseler, H. A., Levine, A. T., & Buckley, D. (2017). Peripheral visual reaction time is faster in deaf adults and British sign language interpreters than in hearing adults. *Frontiers in Psychology*, 8(50), 1–10. <https://doi.org/10.3389/fpsyg.2017.00050>
- Crasborn, O., & Sloetjes, H. (2008). *Enhanced ELAN functionality for sign language corpora: Proceedings of the 3rd workshop on the representation and processing of sign languages: Construction and exploitation of sign language corpora*. LREC 2008, Sixth international conference on language resources and evaluation, Marrakech, Morocco. https://www.researchgate.net/publication/215685477_Enhanced_ELAN_functionality_for_sign_language_coropora
- Crasborn, O. A. (2010). What does “Informed consent” mean in the internet age? Publishing sign language corpora as open content. *Sign Language Studies*, 10(2), 276–290. <https://doi.org/10.1353/sls.0.0044>
- Crasborn, O. A. (2014). Transcription and notation methods. In E. Orfanidou, B. Woll, & G. Morgan (Eds.), *Research methods in sign language studies: A practical guide* (Vol. 6, pp. 74–89). Wiley. <https://doi.org/10.1002/9781118346013.ch5>
- Duranti, A. (2006). Transcripts, like shadows on a wall. *Mind, Culture, and Activity*, 13(4), 301–310. https://doi.org/10.1207/s15327884mca1304_3
- Dynel, M. (2014). On the part of ratified participants: Ratified listeners in multi-party interactions. *Brno Studies in English*, 40(1), 27–44. <https://doi.org/10.5817/BSE2014-1-2>
- Egbert, M. M. (1997). Some interactional achievements of other-initiated repair in multiperson conversation. *Journal of Pragmatics*, 27(5), 611–634. [https://doi.org/10.1016/S0378-2166\(96\)00039-2](https://doi.org/10.1016/S0378-2166(96)00039-2)
- Eisner, W. (2001). *Comics & sequential art* (21th ed.). Poorhouse Press.
- Ekberg, S. (2012). Addressing a source of trouble outside of the repair space. *Journal of Pragmatics*, 44(4), 374–386. <https://doi.org/10.1016/2012.01.006>
- Emmorey, K. (2003). *Perspectives on classifier constructions in sign language: Workshop on classifier constructions*. Lawrence Erlbaum Associates. https://www.researchgate.net/publication/238293697_Perspectives_on_Classifier_Constructions_in_Signed_Languages
- Esmail, J. (2008). *The discourse of embodiment in the nineteenth century British and North American sign language debates* (Publication Number NR69936) [Ph.D.]. Queen's University. ProQuest Dissertations & Theses Global. Ann Arbor, Canada. https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR69936.PDF?oclc_number=794618443
- Ferrara, L., & Johnston, T. (2014). Elaborating who's what: A study of constructed action and clause structure in Auslan (Australian sign language). *Australian Journal of Linguistics*, 34(2), 193–215. <https://doi.org/10.1080/07268602.2014.887405>
- Foucault, M. (1976). Ceci n'est pas une pipe [This is not a pipe]. *October*, 1, 6–21. <https://doi.org/10.2307/778503>
- Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gathering*. The Free Press.
- Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology*, 19(1), 283–307. <https://doi.org/10.1146/annurev.an.19.100190.001435>
- Goodwin, M. H., & Goodwin, C. (2012). Car talk: Integrating texts, bodies, and changing landscapes. *Semiotica*, 2012(191), 257–286. <https://doi.org/10.1515/sem-2012-0063>
- Greer, T. (2015). Appealing to a broker: Initiating third-person repair in mundane second language interaction. *Novitas-ROYAL (Research on Youth and Language)*, 9(1), 1–14. <https://files.eric.ed.gov/fulltext/EJ1167203.pdf>
- Groeber, S., & Pochon-Berger, E. (2014). Turns and turn-taking in sign language interaction: A study of turn-final holds. *Journal of Pragmatics*, 65, 121. <https://doi.org/10.1016/j.pragma.2013.08.012>
- Gustines, G. G. (2014, January 9). One of Marvel's avengers turns to sign language. *ArtsBeat, New York Times Blog*. https://artsbeat.blogs.nytimes.com/2014/07/24/one-of-marvels-avengers-turns-to-sign-language/?_php=true&_type=blogs&_php=true&_type=blogs&_r=1&
- Haddington, P., & Rauniomaa, M. (2014). Interaction between road users: Offering space in traffic. *Space and Culture*, 17(2), 176–190. <https://doi.org/10.1177/1206331213508498>
- Heath, C., & Luff, P. (2012a). Embodied action and organizational activity. In T. S. Jack Sidnell (Ed.), *The handbook of conversation analysis* (pp. 281–307). John Wiley & Sons. <https://doi.org/10.1002/9781118325001.ch14>

- Heath, C., & Luff, P. (2012b). Some 'technical challenges' of video analysis: Social actions, objects, material realities and the problems of perspective. *Qualitative Research*, 12(3), 255–279. <https://doi.org/10.1177/1468794112436655>
- Heath, C., Luff, P., & Hindmarsh, J. (2010). *Video in qualitative research: Analysing social interaction in everyday life*. SAGE.
- Hepburn, A., & Bolden, G. B. (2012). The conversation analytic approach to transcription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 57–76). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118325001.ch4>
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299–345). Cambridge University Press. https://www.researchgate.net/publication/244414507_A_Change-of-State-Token_and_Aspects_of_its_Sequential_Placement
- Hitzel, E. (2015). Introduction. In *Effects of peripheral vision on eye movements: A virtual reality study on gaze allocation in naturalistic tasks* (pp. 5–44). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08466-0_1
- Hjulstad, J. (2017). *Embodied participation: In the semiotic ecology of a visually-oriented virtual classroom* Norwegian University of Science and Technology, Faculty of Humanities. <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2434026>
- Hoffmann-Dilloway, E. (2011). Writing the smile: Language ideologies in, and through, sign language scripts. *Language and Communication*, 31(4), 345–355. <https://doi.org/10.1016/j.langcom.2011.05.008>
- Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Polity Press. <https://doi.org/10.1017/S0047404500233046>
- Ince, I. F., & Kim, J. W. (2011). A 2D eye gaze estimation system with low-resolution webcam images. *EURASIP Journal on Advances in Signal Processing*, 2011(1), 40. <https://doi.org/10.1186/1687-6180-2011-40>
- Ivarsson, J. (2010). Developing the construction sight: Architectural education and technological change. *Visual Communication (London, England)*, 9(2), 171–191. <https://doi.org/10.1177/1470357210369883>
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). John Benjamins. <http://liso-archives.liso.ucsb.edu/Jefferson/Transcript.pdf>
- Johnson, K. L. (1991). Miscommunication in interpreted classroom interaction. *Sign Language Studies*, 1070(1), 1–34. <https://doi.org/10.1353/sls.1991.0005>
- Kaneko, M., & Mesch, J. (2013). Eye gaze in creative sign language. *Sign Language Studies*, 13(3), 372–400. <https://doi.org/10.1353/sls.2013.0008>
- Kauling, E. J. (2012). *Look who's talking. Turn taking in sign language interpreted interaction*. University of Amsterdam. https://www.academia.edu/12780230/Look_whos_talking_Turntaking_in_sign_language_interpreted_interaction
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26(C), 22–63. [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4)
- Kensy, C., Natasha, A., & Susan, G.-M. (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Frontiers in Communication*, 3, 1–23. <https://doi.org/10.3389/fcomm.2018.00023>
- Kitzinger, C. (2012). Repair. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 229–256). Wiley-Blackwell. <https://ebookcentral-proquest-com.ezproxy.oslomet.no>
- Klinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, 100(1), 78–100. <https://doi.org/10.1037/0033-2909.100.1.78>
- Koivisto, A. (2015). Displaying now-understanding: The Finnish change-of-state token aa. *Discourse Processes*, 52(2), 111–148. <https://doi.org/10.1080/0163853X.2014.914357>
- Kuttner, P. J., Weaver-Hightower, M. B., & Sousanis, N. (2020). Comics-based research: The affordances of comics for research across disciplines. *Qualitative Research*, 21(2), 195–214. <https://doi.org/10.1177/1468794120918845>
- Laurier, E. (2013). Before, in and after: Cars making their way through roundabouts. In P. Haddington, L. Mondada, & M. Neville (Eds.), *Interaction and mobility* (Vol. 20, pp. 210–242). De Gruyter. <https://doi.org/10.1515/9783110291278.210>
- Laurier, E. (2014). The graphic transcript: Poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. *Geography Compass*, 8(4), 235–248. <https://doi.org/10.1111/gec3.12123>
- Laurier, E. (2019). The panel show: Further experiments with graphic transcripts and vignettes. *Social Interaction. Video-Based Studies of Human Sociality*, 2(1). <https://doi.org/10.7146/si.v2i1.113968>
- Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society*, 32(2), 177–201. <https://doi.org/10.1017/S004740450332202X>
- Lewandowska-Tomaszczyk, B. (2017). Partial perception and approximate understanding. *Research in Language*, 15(2), 129–152. <https://doi.org/10.1515/rela-2017-0009>
- Linell, P. (1982). *The written language bias in linguistics*. Linköping University Electronic Press. <https://doi.org/10.4324/9780203342763>
- Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations* (2nd ed.). Routledge. <https://ebookcentral-proquest-com.ezproxy.oslomet.no/lib/hioa/detail.action?docID=199577&pq-origsite=primo>
- Manrique, E. (2016). Other-initiated repair in Argentine sign language. *Open Linguistics*, 2(1), 35. <https://doi.org/10.1515/opli-2016-0001>

- Manrique, E. (2017). *Achieving mutual understanding in Argentine sign language (LSA)*. Radboud University Nijmegen. https://pure.mpg.de/rest/items/item_2468462_11/component/file_2468466/content
- Manrique, E., & Enfield, N. J. (2015). Suspending the next turn as a form of repair initiation: Evidence from Argentine sign language. *Frontiers in Psychology*, 6, 1–21. <https://doi.org/10.3389/fpsyg.2015.01326>
- Marstrand, A. K., & Svennevig, J. (2018). A preference for non-invasive touch in caregiving contexts. *Social Interaction. Video-Based Studies of Human Sociality*, 1(2). <https://doi.org/10.7146/si.v1i2.110019>
- McCloud, S., Lappan, B., & Martin, M. (1994). *Understanding comics*. HarperPerennial.
- McIlvenny, P. (2014). Vélomobile formations-in-action: Biking and talking together. *Space and Culture*, 17(2), 137–156. <https://doi.org/10.1177/1206331213508494>
- McKee, R. L. (2011). So, well, whatever: Discourse functions of palm-up in New Zealand sign language. *Sign Language & Linguistics*, 14(2), 213–247. <https://doi.org/10.1075/sll.14.2.01mck>
- Mondada, L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, 43(2), 542–552. <https://doi.org/10.1016/j.pragma.2010.08.019>
- Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2019). Transcribing silent actions: Multimodal approach of sequence organization. *Social Interaction. Video-Based Studies of Human Sociality*, 2(1). <https://doi.org/10.7146/si.v2i1.113150>
- Napier, J. (2007). Cooperation in interpreter-mediated monologic talk. *Discourse & Communication*, 1(4), 407–432. <https://doi.org/10.1177/1750481307082206>
- Nikander, P. (2008). Working with transcripts and translated data. *Qualitative Research in Psychology*, 5(3), 225–231. <https://doi.org/10.1080/14780880802314346>
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43–72). Academic Press. <http://www.sscnet.ucla.edu/anthro/faculty/ochs/articles/ochs1979.pdf>
- Petitta, G., Di Renzo, A., Chiari, I., & Rossini, P. (2013). Sign language representation: New approaches to the study of Italian sign language (LIS). In G. Petitta, A. Di Renzo, I. Chiari, P. Rossini, L. Meurant, A. Sinte, M. Van Herreweghe, & M. Vermeerbergen (Eds.), *Sign language research, uses and practices* (1st ed., pp. 137–158). De Gruyter. <http://www.jstor.org/stable/j.ctvbk4dr.9>
- Pizzuto, E. A., Chiari, I., & Rossini, P. (2011). Representing signed languages: Theoretical, methodological and practical issues. In Fabrizio Serra (Ed.) *Rivista di psicolinguistica applicata*: XI, 3 (pp. 205–241). Pisa. <http://digital.casalini.it/2496458>
- Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica*, 78(1–2), 75–100. <https://doi.org/10.1515/semi.1990.78.1-2.75>
- Rosenthal, A. (2009). Lost in transcription: The problematics of commensurability in academic representations of American sign language. *Text & Talk*, 29(5), 595. <https://doi.org/10.1515/TEXT.2009.031>
- Sacks, H., Jefferson, G., & Schegloff, E. A. (1995). *Lectures on conversation* (One paperback vol ed.). Blackwell. <https://doi.org/10.1002/9781444328301>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn taking for conversation. *Language*, 50(4), 696–735. <https://doi.org/10.1016/B978-0-12-623550-0.50008-2>
- Sallandre, M.-A., & Garcia, B. (2013). Epistemological issues in the semiological model for the annotation of sign languages. In L. Meurant, A. Sinte, M. Van Herreweghe, M. Vermeerbergen, & A. Sinte (Eds.), *Sign language research, uses and practices: Crossing views on theoretical and applied sign language linguistics* (pp. 159–178). De Gruyter. <http://ebookcentral.proquest.com/lib/hioa/detail.action?docID=893554>
- Sallandre, M.-A., & Garcia, B. (2020). Contribution of the semiological approach to deixis-anaphora in sign language: The key role of eye-gaze. *Frontiers in Psychology*, 11, 583763. <https://doi.org/10.3389/fpsyg.2020.583763>
- Sanford, A. J., & Moxey, L. M. (1995). Aspects of coherence in written language: A psychological perspective. In M. A. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (Vol. 31, pp. 161–187). J. Benjamins.
- Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 97(5), 1295–1345. <https://doi.org/10.1086/229903>
- Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, 23(3), 499–546. <https://doi.org/10.1080/01638539709545001>
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53(2), 361–382. <https://doi.org/10.1353/lan.1977.0041>
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U., Schlobinski, P., & Uhmann, S. (2011). A transcription system for conversation analysis. *Linguistische Berichte*, 12, 1–51. <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2011/px-gat2-englisch.pdf>
- Sheyhshhe, M. A. (2013). *Native Americans in comic books: A critical study*. McFarland, Incorporated, Publishers. <https://books.google.no/books?id=Nf5tAAAAQBAJ>
- Skedsmo, K. (2020a). Multiple other-initiations of repair in Norwegian sign language. *Open Linguistics*, 6(1), 532–566. <https://doi.org/10.1515/opli-2020-0030>

- Skedsmo, K. (2020b). Other-initiations of repair in Norwegian sign language. *Social Interaction. Video-Based Studies of Human Sociality*, 3(2). <https://doi.org/10.7146/si.v3i2.117723>
- Stone, C., & West, D. (2012). Translation, representation and the Deaf 'voice'. *Qualitative Research*, 12(6), 645–665. <https://doi.org/10.1177/1468794111433087>
- Supalla, S. J., Cripps, J. H., & Byrne, A. P. J. (2017). Why American sign language gloss must matter. *American Annals of the Deaf*, 161(5), 540–551. <https://doi.org/10.1353/aad.2017.0004>
- Swisher, M., Christie, K., & Miller, S. (1989). The reception of signs in peripheral vision. *Sign Language Studies*, 1063(1), 99–125. <https://doi.org/10.1353/sls.1989.0011>
- Takkinen, R. (2005). Some observations on the use of HamNoSys (Hamburg notation system for sign languages) in the context of the phonetic transcription of childrens signing. *Sign Language & Linguistics*, 8(1–2), 99–118. <https://doi.org/10.1075/sll.8.1.05tak>
- Ten Have, P. (2007). *Doing conversation analysis* (2nd ed.). SAGE Publications, Ltd. <https://doi.org/10.4135/9781849208895>
- Thompson, R., Kluender, R., & Emmorey, K. (2006). *Eye gaze in American sign language: Linguistic functions for verbs and pronouns* [Ph.D]. University of California. ProQuest Dissertations Publishing.
- Todorović, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research*, 46(21), 3549–3562. <https://doi.org/10.1016/j.visres.2006.04.011>
- Valli, C., Lucas, C., Mulrooney, K. J., & Villanueva, M. (2011). *Linguistics of American sign language: An introduction* (5th ed.). Gallaudet University Press.
- Van Herreweghe, M. (2002). Turn-taking mechanisms and active participation in meetings with Deaf and hearing participants in Flanders. In C. Lucas (Ed.), *Turn-taking, fingerspelling and contact in signed languages* (Vol. 8, pp. 73–103). Gallaudet University Press. <http://hdl.handle.net/1854/LU-687117>
- Wallner, L. (2017a). *Framing education: Doing comics literacy in the classroom* Linköping University. DiVA. Linköping University Electronic Press. <http://liu.diva-portal.org/smash/get/diva2:1150907/FULLTEXT03.pdf>
- Wallner, L. (2017b). Speak of the bubble - constructing comic book bubbles as literary devices in a primary school classroom. *Journal of Graphic Novels & Comics*, 8(2), 173–192. <https://doi.org/10.1080/21504857.2016.1270221>
- Wallner, L. (2018). Gutter talk: Co-constructing narratives using comics in the classroom. *Scandinavian Journal of Educational Research*, 63(6), 819–838. <https://doi.org/10.1080/00313831.2018.1452290>
- Wilkes-Gibb, D. (1995). Coherence in collaboration. In M. A. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (Vol. 31, pp. 239–268). J. Benjamins. <https://doi.org/10.1075/tsl.31.09wil>
- Willemsen, A., Gosen, M., Koole, T., & De Gloppe, K. (2020). Gesture, gaze and laughter. *Social Interaction. Video-Based Studies of Human Sociality*, 2(2). <https://doi.org/10.7146/si.v2i2.118045>
- Wilson, H. R., Wilkinson, F., Lin, L.-M., & Castillo, M. (2000). Perception of head orientation. *Vision Research*, 40(5), 459–472. [https://doi.org/10.1016/S0042-6989\(99\)00195-9](https://doi.org/10.1016/S0042-6989(99)00195-9)