



This is a repository copy of *Talking in Time: the development of a self-administered Conversation Analysis based training programme for cochlear implant users.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/146532/>

Version: Published Version

---

**Article:**

Wells, W., Beeston, A., Bradley, E. et al. (3 more authors) (2019) Talking in Time: the development of a self-administered Conversation Analysis based training programme for cochlear implant users. *Cochlear Implants International*, 20 (5). pp. 255-265. ISSN 1467-0100

<https://doi.org/10.1080/14670100.2019.1625185>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>



# Cochlear Implants International

## An Interdisciplinary Journal

ISSN: 1467-0100 (Print) 1754-7628 (Online) Journal homepage: <https://www.tandfonline.com/loi/ycii20>

## *Talking in Time*: The development of a self-administered conversation analysis based training programme for cochlear implant users

Bill Wells, Amy V. Beeston, Erica Bradley, Guy J. Brown, Harriet Crook & Emina Kurtić

To cite this article: Bill Wells, Amy V. Beeston, Erica Bradley, Guy J. Brown, Harriet Crook & Emina Kurtić (2019) *Talking in Time*: The development of a self-administered conversation analysis based training programme for cochlear implant users, *Cochlear Implants International*, 20:5, 255-265, DOI: [10.1080/14670100.2019.1625185](https://doi.org/10.1080/14670100.2019.1625185)

To link to this article: <https://doi.org/10.1080/14670100.2019.1625185>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 24 Jun 2019.



[Submit your article to this journal](#)



Article views: 207



[View Crossmark data](#)

# Talking in Time: The development of a self-administered conversation analysis based training programme for cochlear implant users

Bill Wells<sup>1</sup>, Amy V. Beeston <sup>2</sup>, Erica Bradley<sup>3</sup>, Guy J. Brown <sup>2</sup>, Harriet Crook<sup>3</sup>, Emina Kurtić<sup>1</sup>

<sup>1</sup>Department of Human Communication Sciences, University of Sheffield, Sheffield, UK, <sup>2</sup>Department of Computer Science, University of Sheffield, Sheffield, UK, <sup>3</sup>Department of Neurotology, Sheffield Teaching Hospitals NHS Trust, Sheffield, UK

**Objectives:** Training software to facilitate participation in conversations where overlapping talk is common was to be developed with the involvement of Cochlear implant (CI) users.

**Methods:** Examples of common types of overlap were extracted from a recorded corpus of 3.5 hours of British English conversation. In eight meetings, an expert panel of five CI users tried out ideas for a computer-based training programme addressing difficulties in turn-taking.

**Results:** Based on feedback from the panel, a training programme was devised. The first module consists of introductory videos. The three remaining modules, implemented in interactive software, focus on non-overlapped turn-taking, competitive overlaps and accidental overlaps.

**Discussion:** The development process is considered in light of feedback from panel members and from an end of project dissemination event. Benefits, limitations and challenges of the present approach to user involvement and to the design of self-administered communication training programmes are discussed.

**Conclusion:** The project was characterized by two innovative features: the involvement of service users not only at its outset and conclusion but throughout its course; and the exclusive use of naturally occurring conversational speech in the training programme. While both present practical challenges, the project has demonstrated the potential for ecologically valid speech rehabilitation training.

**Keywords:** Conversation analysis, Overlapping talk, Cochlear implant users, Rehabilitation software, Self-administered training, User involvement

## Introduction

Loss of hearing is not simply the absence of sound; it results in a more limited capacity to take part in social life, which is primarily enacted through conversation. Inability to fully participate in conversation can lead to being treated differently. One reason why listeners with hearing impairments, including cochlear implant users, may find it hard to take part in a conversation is that when participants in the conversation take turns to speak, the exchange of turns happens rapidly (Levinson and Torreira, 2015). Participants need to attend to the various cues that speakers use to signal that they are continuing to speak and then to project that they are about to finish. A participant in the role of listener may sometimes miss some of these cues. One consequence is that this participant

may miss the chance to take a turn when the speaker has finished. Another possibility is that the participant may inadvertently take a turn before the current speaker has finished, resulting in the situation where the two participants are speaking in overlap, in which case the current speaker might react to the incoming talk as an unwarranted interruption (Tye-Murray and Witt, 1996).

More generally, the occurrence of overlapping talk, where two or more conversational participants are speaking simultaneously, presents a major challenge for conversational participants with hearing impairments as well as for their hearing conversational partners (Skelt, 2013). Over and above the social awkwardness that might arise from the scenario just described, there is the challenge, in conversations involving three or more participants, of following what is being said when two or more of the other participants are speaking in overlap – a situation that is

Correspondence to: Guy Brown, Department of Computer Science, University of Sheffield, Sheffield, UK. Email: g.j.brown@sheffield.ac.uk

surprisingly common. The aim of the project described in this article was to develop ways to improve cochlear implant users' experience during conversation, by devising training software focussed on activities related to turn-taking and overlap. Ideally these activities would enable the user to practise some strategies for dealing with such problems in daily life.

### *Types of overlapping talk*

In the face-to-face multi-party conversations among four British adults used in the present study (see 'Material' below), 41% of speaker turns are overlapped by another speaker, occupying 16% of the total talking time (cf. Heldner and Edlund (2010) and Kurtić *et al.* (2013) for comparable statistics in other corpora). The frequent occurrence of overlap does not, however, mean that turn-taking is disorganized or random: when a speaker starts talking in overlap it is often intentional. While the possible reasons for overlapping are various (Tannen, 1983) two basic motives can be identified: either to be *collaborative* with current speaker, showing solidarity with what is being said; or to *compete* with the current speaker for occupancy of the floor. In the *competitive* scenario, the overlapper demonstrates the desire to prevent the current speaker from finishing what he wants to say, by taking over the floor immediately; so the overlapper usually starts talking at an early point in the current speaker's turn and continues talking till the overlappee drops out or fights back. The overlapper often raises the pitch and volume of her speech above its usual conversational level. The combination of fundamental frequency (F0 – the main acoustic correlate of pitch) and intensity (the main acoustic correlate of loudness) is the most prominently used phonetic resource for turn competition (Kurtić *et al.*, 2013). This behaviour can be described as an interruption, though not all interruptions occur in overlap (Schegloff, 2001). In the recorded corpus used in the present research, 28% of overlaps were identified as competitive.

Even though interrupting may sometimes be evaluated negatively by observers (Hilton, 2016), it is nevertheless useful to have the ability to overlap in a competitive way in order to take the floor. We may want to do this if the current speaker is dominating the conversation so that the other participants are not getting the chance to say anything; or if we disagree with what the current speaker is saying; or if we want to change the topic of the conversation. Overlapping competitively is thus a way in which one can assert one's rights and responsibilities as a participant in a conversational situation:

Finding room to talk is up to speakers. Indeed, it is incumbent upon speakers, if they are observing this system, to find things to say and in the

conversation to say them. A person who gives up after a single try is perceived by overlap-favouring speakers as being uncooperative, withholding, even sulking. (Tannen, 1983: 125)

Instead of competing for the floor, a participant may overlap in order to display a *collaborative* stance towards the current speaker. Most collaborative overlaps are quite different from competitive ones. They often occur late in the current speaker's turn, are brief and the overlapper usually uses low pitch and volume (Kurtić *et al.*, 2013). It is useful in conversation to be able to identify when another speaker's overlaps are collaborative. It is also useful to be able to overlap collaboratively, since this is a way to show that we are following what the current speaker is saying and thus engaging with the talk (Tannen, 1983; Hilton, 2016).

In both competitive and collaborative overlaps, the speaker's incoming talk in overlap is intentional. However, overlaps can also happen by accident, as a result of the latitude with regard to selection of the next speaker that is inherent in the turn-exchange system (Sacks *et al.*, 1974; Levinson and Torreira, 2015). Accidental overlaps typically happen when two or more participants start to talk together following the (apparent) end of the current speaker's turn, i.e. at a turn transition relevance place. Usually, one speaker drops out immediately on realizing that they are in overlap, leaving the floor clear for the other speaker (Jefferson, 1984, 1987; Schegloff, 2000; see Kurtić, 2011 for review).

The key participant behaviours involved in overlap can be summarized as follows.

*Temporal organization of the turn.* Important features for displaying an overlap as competitive include: the early placement of the incoming talk in relation to the turn in progress; the recycling (repetition) of turn beginnings; overlapped speaker cutting off the turn in progress (Kurtić *et al.*, 2013).

*Prosodic design of the turn.* For turn competition, important features include raised loudness and pitch height (French and Local, 1983) and speech rate (Kurtić *et al.* 2010); whereas for non-competitive incomings these features are likely to be absent; instead the incoming speaker is likely to match the pitch contour of the current speaker's turn (Kurtić and Gorisch, 2018).

*Nonverbal design of the turn.* It has been proposed that gestures and gaze are relevant resources for overlap management in face-to-face discourse. Lee *et al.* (2008) found that hand movements helped to discriminate between turn-competitive and non-competitive overlaps in a corpus of acted dialogues. In a study of French natural conversations, Mondada and Oloff (2011) showed that continuing vs. abandoning

gesturing during overlap is associated with how problematic participants take the overlap to be. With regard to gaze, Auer (2017) suggests that the current speaker may use gaze to invite one of several potential next speakers to take the next turn in overlap.

### *Competence in managing overlapping talk*

From the preceding section, it is evident that one facet of a conversational participant's communicative competence is competence with overlap and that being overlap-competent involves a range of skills. As a *speaker*, a conversational participant needs to be able to achieve the functionally distinct types of overlapping talk just described, for example:

- a) Come in at the right place and in the right way in the normal exchange of turns;
- b) Drop out if one finds oneself starting up simultaneously with another speaker;
- c) Chime in appropriately in collective greetings, toasts etc;
- d) Provide feedback to a speaker who has taken the floor e.g. to tell a story;
- e) Interrupt a speaker;
- f) Hold onto the floor when interrupted if one does not wish to surrender the floor.

As a *listener*, the participant needs to be able to follow what is happening when overlaps occur in the conversation:

- a) Identify and track what another speaker is *doing* when in overlap, i.e. the speaker's social action. For example, is she completing the speaker's turn in a collaborative way; giving supportive feedback; joining in a collective activity; or interrupting the speaker?
- b) Identify and track what other speakers are *saying* when in overlap. Decoding and parsing the words produced by speakers in overlap is often difficult even for normal hearing listeners. Where it is possible, it will help the listener (a) to keep track of the topic of the talk and (b) to identify the speaker's social action or intent (as above).
- c) Identify and track the schisming of a wider multi-party conversation into sub-conversations, and their reforming into a plenary conversation (Egbert, 1997).

The overlap-competent participant will thus be skilled in interpreting and reacting to functionally distinct types of overlapping talk when used by other participants in the conversation.

### *Cochlear implants and overlapping talk*

It is important to consider the particular challenges posed by overlapping talk for a cochlear implant (CI) user. These challenges are threefold: first, to identify who is speaking; second, to identify what the overlapping speakers are saying, i.e. to decode and parse the words; and third, to identify what the

overlapping speakers are doing, i.e. to identify the social action or actions that they are trying to accomplish.

In considering these three challenges, account has to be taken of the nature of the speech signal, the way in which these signals are processed by the cochlear implant and the resulting speech information that is transmitted to the CI user. Because the implant typically has just 22 electrodes, as opposed to 3500 inner hair cells in a healthy cochlea, the number of frequency channels available is small and spectral resolution is low. Harmonics are not well represented as it is the spectral envelope of the speech that is encoded, resulting in loss of the temporal fine structure of the signal and poor representation of the fundamental frequency (F0) (Peng *et al.*, 2008; Van De Velde *et al.*, 2015). As noted above, F0 is the main acoustic correlate of the perceived pitch for unimpaired listeners. However, most CI listeners report changes in F0 as timbre, rather than pitch *per se* (Crew *et al.*, 2016). Similarly, the ability to use spatial location information (which is an important cue for segregating sound sources) is impaired in CI users because interaural time and intensity differences are poorly represented (Litovsky *et al.*, 2009). CI users thus need to make sense of complex mixtures of sound without the benefit of cues from pitch or location. This is most likely due to the emphasis in CI design on signal processing strategies that help with the identification of spectral (consonant and vowel) features. One consequence of these characteristics of the CI-mediated hearing of speech is that the user may have difficulty in identifying the gender of speakers (Kovačić and Balaban, 2009), which makes it harder to distinguish one talker from another and thus to identify who is speaking. A further consequence is that even after cochlear implantation, listeners have considerable difficulty with the perception of the pitch patterns of speech that are used in tone, stress and intonation systems. For children growing up as CI users this results in delay and difficulty with producing the intonation of the language being acquired (Most and Peled, 2007; O'Halpin, 2010; Peng *et al.*, 2009; Snow and Ertmer, 2012).

These characteristics of the CI signal will impact on the user's ability to deal with overlap. The talk of overlapping speakers can be viewed as a noise source that masks the signal from the 'target' speaker, that is the speaker whom the listener is trying to identify and then attend to in order to track what is being said. This is a well-known problem that confronts hearing as well as hearing-impaired listeners, though is particularly challenging for the latter (Fuller *et al.*, 2014). Furthermore, in conversational interactions, the listener needs to be able to identify and track what other speakers are *doing* when in overlap, i.e.



what is the speaker's social action. Since, as explained above, in addition to the speaker's words and their meaning, there are temporal and prosodic cues to social action that are present in the overlapping speaker's turn, indicating whether or not the overlap is competing for the floor, the relative lack of access for the CI user to such temporal and prosodic features is likely to impair the CI user's ability to track the social actions embodied in overlapping talk.

One behavioural consequence may be to withdraw from conversation where overlap is prevalent. It may have other consequences too. Thus Tye-Murray and Witt (1996) reported that in dialogues between adult CI users and unfamiliar hearing conversational partners, the CI users made significantly more interruptions than their partners, which led the authors to suggest that some of the CI users dominated the conversation with their hearing partners.

The task of following talk in overlap is thus particularly challenging for participants who have a hearing impairment. Speech and language therapists and other professionals may have steered clear of advising CI users about how to deal with situations of overlapping talk, on the basis that it would be just too hard to handle. It has been accepted that even in one-to-one settings, many people who are hearing-impaired, including CI users, need optimum conditions in order to hold a conversation, e.g. quiet background, communication awareness of both participants that they should avoid talking at the same time. However, recent developments in CIs mean that it is now more realistic for users to engage with confidence in conversation where overlapping talk occurs (Luo *et al.*, 2014).

The results of the studies just reviewed imply that when addressing issues for CI users that arise from overlapping talk, the focus should not be exclusively on the limitations of, and possible improvements to the CI device itself. Rather, it will be helpful to focus in addition on interactional functions, so that the CI user can understand what is potentially going on when speakers talk in overlap. Users can then be advised on how to participate in conversation at points where overlap occurs and is interactionally relevant. Conversation analytic research focusing on overlapping talk in interactions between a young cochlear implant user and his mother underlines the crucial role of the co-participant(s) in maximizing the CI user's full participation in the cut and thrust of conversation (Anstey and Wells, 2013). This suggests the inclusion of CI users and their regular communication partners in the implementation of training materials.

### **The development of *Talking in Time***

The research described in the previous section suggests that there is a *prima facie* case for developing training

materials to assist CI users who are keen to develop skills in conversational turn-taking, including the management of overlapping talk. The clinical members of our team followed this up in the context of a small-scale audit of adult CI users in South Yorkshire using the Hearing Implant Sound Quality questionnaire (HISQUI) (Amann and Anderson, 2014). From this, more complex aspects of communication skills were identified which respondents felt could benefit from further rehabilitation, with a view to improving quality of life and participation in employment as well as social activity. One of these aspects was simultaneous or overlapping talk.

*Talking in Time*, the project described here, had as its main aim to develop rehabilitative software to assist CI users to participate in multi-speaker conversations where overlapping talk is common. When considering how to approach the development of a rehabilitation programme, the team was aware of the need for a cost-effective solution. New technology has opened up options for access to self-administered communication-related training that can be carried out at home, such as the *SWORD* program for people with apraxia of speech resulting from a stroke (Varley *et al.*, 2016). There have also been recent developments in computer-based learning that focus directly on conversational interaction rather than isolated linguistic skills, such as *Better Conversations with Aphasia* (Beeke *et al.*, 2013), in which, following the methodological principles of Conversation Analysis research, the video material used is drawn entirely from recordings of naturalistic interactions rather than staged or scripted dialogues. In addition to drawing on the strengths of programmes such as *SWORD* and *Better Conversations with Aphasia*, the research team was committed to involving CI users not just in the trialling of the eventual rehabilitation product but also in the actual process of developing the software, in order to ensure that it would address the needs of users.

### **Material**

One tenet of Conversation Analysis research is to restrict the data analysed to recordings of naturalistic talk-in-interaction. It is considered desirable where possible, to restrict the extracts used for pedagogical purposes to that source too. With this in mind, in *Talking in Time* the user works with conversational extracts selected from naturalistic recordings. The recordings, which had been made at the University of Sheffield in order to create a British English corpus for an earlier project on overlapping talk, are of unscripted face-to-face conversations between four young adult friends seated round a table. Individual headset microphones were used to record the audio signal onto separate channels for each speaker,

making it possible to analyse instances of overlapping talk in detail. Video recordings were made using two camera angles. The recordings had been transcribed orthographically, then segmented into turns. All instances of overlap were identified and classified. The conversations contained some portions that could not be used for confidentiality reasons so these portions were removed from the corpus. All annotation was carried out using the ELAN program (Version 4.6.2) (Wittenburg *et al.*, 2006). Ethics approval was obtained from the University of Sheffield. Consent for use of the recordings was obtained from the four participants prior to the recording. Details of the recordings, transcription and annotation can be found in Kurtić *et al.* (2012).

#### *The project team and the expert CI user panel*

The project team consisted of a speech and language therapist (Bradley) and an audiological scientist (Crook) working clinically with CI users, two computer scientists with expertise in speech processing and software development (Beeston, Brown) and two linguists specializing in the clinical application of conversation analysis and phonetics (Kurtić, Wells). Involving patients in service development and research is a priority for the UK National Health Service and therefore for collaborating universities, in order to ensure that what is developed will be relevant and is what patients want and will use. To achieve this, an expert panel of five adult CI users of varying age, gender and hearing history were recruited. CI users are a varied population, including those deaf since birth and those with acquired deafness, with different experiences, strategies and expectations. It was therefore deemed important to involve a range of users in the development process. An important role of the speech and language therapist and the audiological scientist in the project was to facilitate this recruitment process, thereby bringing together those developing the software and those who will use it.

#### *Panel meetings*

The panel participated in eight meetings with the project team, spread over the duration of the project (12 months). Their main role was to contribute to the development of the computer-based, self-administered training programme that came to be known as *Talking in Time*. While the main focus of *Talking in Time* in its present form is on the development of awareness and listening skills, the opportunity was taken also to try out some speaking exercises with the view to their incorporation in a later version of the software.

There was no attrition, each meeting being attended by at least four of the five CI users and by at least five of the six researchers. Each session consisted of a mix

of group discussion and individual sessions with a team member, working on pilot exercises presented on a laptop computer. The plenary discussions were audio recorded and summarized in written form after the meeting by one of the project team.

There were various iterations of material selection and task development over the course of the meetings, which allowed the team to home in on major themes of difficulty, using a variety of means. These included:

- a) PowerPoint presentation (Version 14.0) ('PowerPoint,' 2011) followed by general discussion.
- b) PowerPoint presentation and clicker key-press responses; this enabled the team to gather instant but anonymous feedback on the materials being tried out.
- c) psychoPy presentation (Peirce, 2014) for prototyping listening tasks: several listening tasks intended for the software were presented as psychoacoustic experiments and tested out individually by the panel members.
- d) iMovie (Version 10.0.3) presentation for simulation of speaking tasks: this allowed the team to present video extracts of conversation and make audio recordings of users' verbal responses to prototype the speaking tasks intended for the software.
- e) paper-based mock-ups to elicit feedback on interface design.

Numerous issues relevant to the development of *Talking in Time* were raised at the panel meetings. Among the most notable were:

- a) Comments on the difficulty of following the conversation extracts from the corpus. One reason was the lack of explicit contextual information. This issue was subsequently addressed in the software by providing a written summary of the topic being discussed in each extract on the screen (see next section). Another reason was the very informal nature of the talk, resulting in fast speech rate and abundance of connected speech processes. This was addressed by providing an orthographic transcription of the extract on the screen. This difficulty also led to discussion around the observation that even people with unimpaired hearing will struggle to identify every word in an informal conversation and that it may therefore be important to try and extract the gist of what is being said in a turn as a basis for identifying the social action the speaker is trying to accomplish.
- b) Discussion around whether the transcript of the conversation extract should be presented as a subtitle superimposed on the video itself, or in a separate box outside the video frame. It was decided to go with the latter option, as it was the preference of the panel members.
- c) Limitations of the video corpus, particularly regarding the visibility of the speaker's mouth, which is important for lip reading. This could not be addressed as the corpus had been collected and transcribed for an earlier project and there was no

funding available in the present project for a new video corpus.

- d) Resistance to the premise of being assisted to practise incomings in overlap, based on the belief that it is rude to interrupt. This led to discussions about being assertive vs. being rude in conversation. In terms of the software, it made the team aware of the need to address this issue explicitly in the introduction to the *Talking in Time* programme (see description of Module 1 below).

Panel members reported that they found attendance at the meetings worthwhile. They felt that their understanding of conversation had improved and they enjoyed the opportunity to meet other CI users. The members of the research team also felt they had benefited greatly from participating: the health service members in terms of increased knowledge about conversation analysis and the university-based members in terms of increased understanding of the communicative life of CI users as well as the role of health professionals in this regard.

### *Dissemination event*

At the end of the project, an event was held at the University of Sheffield to present *Talking in Time*. It was attended by CI users, family members, speech and language therapists, audiologists as well as researchers and students. The expert panel members contributed to the organization and delivery of this event. In particular, their input ensured that the presentations were accessible to all members of the audience, by means of the following provisions:

- a) Two screens displayed the presentation slides while one further screen showed with a close up of the presenter's face, to enable lip reading;
- b) there was no light behind the presenter, as this could have cast their face in shadow;
- c) there was good lighting on the face of the presenter;
- d) presenters were requested to face the audience and camera throughout their presentation;
- e) presenters were asked to monitor the rate and clarity of speech;
- f) a sound system 'loop enabled' for hearing aids was used;
- g) presenters were asked to provide written information on their slides to supplement the spoken content of the presentation;
- h) audience questions were transcribed in real time by a professional typist and projected on the two presentation screens so that the question was immediately available to the audience in written form.

Members of the audience, including some who did not have a hearing loss, commented afterwards on how easy it had been to follow the presentations and discussion. According to one audience member, 'it was so much easier to focus on the content of the presentation rather than putting more effort into actually

hearing and listening'. A notable highlight of the launch event was a session where audience members were able to question the expert panel about their experiences of participating in everyday conversations, as well as their experience in working on the project. At the end of the day, audience members were able to try out the *Talking in Time* software.

### **Structure and content of *Talking in Time***

The *Talking in Time* software has been developed using Max (version 6.0.8) ('Max,' 2012). Considerations in choice of language included the ability to build for different platforms (Windows, Mac); the robust handling of video and sound; the need for rapid development given the timescale of the project; and download size.

*Talking in Time* comprises four Modules. While Module 1 comprises a series of short introductory videos described below, Modules 2, 3 and 4 are interactive, each consisting of two phases and following the same pattern. Phase 1 is an *Awareness* phase, where the user can get used to watching videos of people having a conversation, in order to become more aware of matters such as: which participant speaks first; whether or not there is a next speaker; and if so, who the next speaker is; as well as the cues that speakers use to signal the different types of overlap. In Phase 2, the *Listening Phase*, the focus is on listening to speakers as they take turns in the conversation. There is practice in identifying participants taking turns to speak one after another versus participants speaking in overlap. The user also gets practice in identifying different types of overlap: competitive, collaborative or accidental.

The interface is the same for each phase of Modules 2, 3 and 4 (Figure 1). On the left of the screen is the video display, where the four participants are seated around a table. Below the video display are three lines of written information. The top line describes the topic of the selected conversational exchange. The middle line shows the words of the first speaker and the bottom line shows the words of the second speaker. On the right of the screen, the task for the user is presented. In the example shown in Figure 1, the task is to answer the question: 'Are there both male and female voices in this recording?'. Below the question are two clickable buttons: 'yes' and 'no'. At the bottom right of the screen are three buttons. 'Settings' can be used to adjust the volume. 'Replay' is for when the user wants to replay the current extract, while 'next' is used to move on to a new extract.

In all phases of Modules 2, 3 and 4 the user receives feedback on the accuracy of each response and has the opportunity to repeat each task as often as is desired. On each trial, a new recorded fragment of real



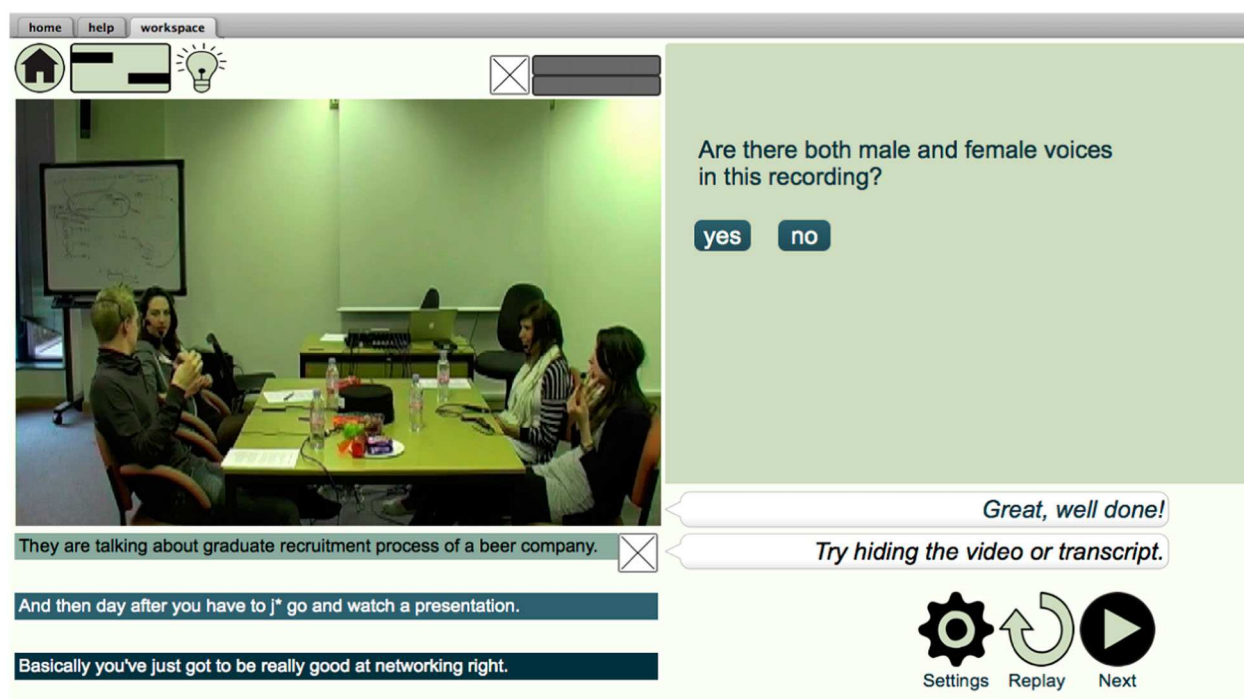


Figure 1. Screenshot of the *Talking in Time* interface.

conversation illustrating the point under consideration is retrieved by the *Talking in Time* programme from its store. In the default presentation of each conversation extract on the screen, the user watches the video of the exchange as well as hearing the audio track. In addition, a transcript of the exchange is visible. In order to make the task more challenging, the user can at any point choose to hide the transcript or the video or both, by clicking on the boxes containing an 'X'. To make listening easier the original sound channel of the video as recorded by the camera was stripped away and replaced by a mono mix of the headset microphone recordings of only those talkers who are involved in the turn-exchange in question. This mono mix was then repeated on both audio channels of the video clip. This removes the room sounds as well as any additional noise that was captured by the camera microphone and substantially improves the sound quality.

Modules 2–4 and their phases, which are described in more detail below, are structured according to a hierarchy of difficulty, based on feedback from the panel of expert CI users. It is therefore envisaged that initially the user will work through the modules and phases in the order in which they are presented. This is not obligatory, however, as the software permits the user to work on phases and modules in any order.

### *Module 1: taking part in conversations*

Module 1 consists of a series of short videos presented by the speech and language therapist on the team, organized into three Phases. Following a general

introduction to the software in Phase 1, Phase 2 provides instructions and guidance on how to use the software interface. Phase 3, 'How conversations work', includes videos that introduce the topic of turn-taking in conversation, the reasons why overlaps happen in conversation and finally, the cues that speakers use to mark the impending end of conversational turns. The videos in Phase 3 could be used independently of the software, by anyone seeking an introduction to turn-taking and overlapping talk. For an example, the reader may access the video clip *Mod1Phase3 TinT*. ([Supplemental material](#)) In this video clip, the analogy of road traffic lights is used to explain how participants can signal the continuation and the impending end of a turn at talk by using grammatical, prosodic and non-verbal cues.

### *Module 2: one speaker at a time*

In Module 2, the focus is on turns that occur *after* the previous speaker has finished, as a preliminary to practice with overlapping talk in Modules 3 and 4. Phase 1 of Module 2 has a preparatory function: to increase the user's general awareness of speakers taking turns in conversation. In Phase 1, the user gets practice in identifying male vs. female speakers in conversation. The rationale is that in a conversation involving more than two people, it may be hard to tell who is speaking. In a mixed conversation, a useful first step is to decide if the speaker is a male or a female. As noted earlier, listeners with hearing difficulties may find this challenging, including those using a cochlear implant for whom the F0 contour may not be well

reproduced. A user who successfully completes this phase should be better able to distinguish between (a) an exchange of conversational turns where the second speaker is of different gender to the first speaker vs. (b) an exchange of conversational turns where the first speaker and the second speaker are of the same gender.

The aim of Phase 2 is to increase awareness of how speakers take turns in conversation, by distinguishing a clear (i.e. non-overlapping) turn taken on time from a clear turn taken late. In conversation it is useful to be able to recognize when a new speaker's turn starts late, since it very often indicates that the new speaker is experiencing some kind of trouble. The trouble may arise for reasons such as not fully hearing the previous speaker's turn; hearing it but not fully understanding it; hearing and understanding it but having some social difficulty with it, e.g. disagreeing with the content or not wanting to accept an invitation that is contained in the prior speaker's turn (Pomerantz and Heritage, 2012). In order to follow what is going on between the participants in a conversation, it is therefore useful to be able to recognize a sign of trouble, such as a delayed start to a turn. In this phase, for each two-turn exchange, the user has to decide if the second speaker's response happens on time or late.

When selecting conversation extracts for the software, the cut off between 'not late' and 'late' was set at 1 s, based on evidence from the research literature (Jefferson, 1989). For pedagogical purposes it was decided to use clear instances of the two categories, as determined by the objective temporal criterion of time measurement supplemented by the researchers' subjective judgements.

### *Module 3: competitive overlaps*

Phase 1, the awareness phase, is designed to further enhance awareness of speakers taking turns in a conversation. In a conversation involving more than two people it can sometimes be hard to tell if a new speaker has started a turn or if the original speaker is continuing to talk. Listeners with hearing difficulties, including those using a cochlear implant, may find this particularly challenging. In this phase, the user gets practice in identifying one speaker vs. multiple speakers. The extracts from the corpus consist either of a single speaker or an exchange involving two speakers. Users who successfully complete this phase will be able to distinguish between (a) an exchange of conversational turns (i.e. where the second speaker is different to the first speaker) vs. (b) a single conversational turn of at least two parts (i.e. produced by one speaker).

Phase 2, the listening phase, focuses specifically on competitive overlaps. In this phase, users get practice

in distinguishing turns that start on time from turns that start early, i.e. in overlap. Users who successfully complete this phase should be able to identify cases where a second speaker's turn overlaps the first speaker's turn and to distinguish these from speaker exchange where there is no overlap. The conversation extracts selected from the corpus for this phase consist of (a) turn exchanges in the clear and (b) overlapping turns where the incoming speaker has been judged to be overlapping in a competitive way. At this point, the reader is advised to view the video clip *DemoModule3.mov*, ([Supplemental material](#)) which demonstrates how the user interacts with the software in general and with Module 3 Phase 2 specifically.

### *Module 4: accidental overlaps*

As explained earlier, a speaker may start in overlap for various reasons, principally (a) to take the floor before the current speaker has finished their turn; (b) to show support for the current speaker in a collaborative way; (c) by accident because it seemed that the current speaker had already finished. Whereas Module 3 focuses exclusively on the first of these possibilities, i.e. competitive overlaps, in Module 4 the user is introduced to the most characteristic feature of accidental overlaps and learns to distinguish them from collaborative and competitive overlaps.

The most characteristic feature of accidental overlap, is that one speaker drops out quickly. In Module 4 Phase 1, the awareness phase, for each example, the user therefore must decide if one speaker drops out quickly (accidental overlaps) or not (collaborative overlaps). In Phase 2 of Module 4 the software selects from its store an accidental, competitive or collaborative overlap and the user has to decide which type of overlap is exemplified.

### *Speaking exercises*

It was not possible within the time and funding available to incorporate speaking exercises into the programme. However, in the future it is planned to include exercises in which the user is prompted to provide different kinds of spoken response to conversational turns presented on the screen. Such exercises will require audio input from the user, i.e. a conversational 'turn' in response to a prompt. The turn will then need to be processed by the software and feedback as to its accuracy will need to be generated. Although these technical issues remain to be addressed, the content of the Speaking phases (Phase 3) for Modules 2, 3 and 4 has been trialled with the user panel, resulting in proposed exercises.

In Phase 3 of Module 2, the user will practise responding to conversational extracts from the corpus. For each trial, after the first turn-constructive unit (TCU) of the speaker's turn in the recorded

extract, the video is muted. The user then has to take the next turn ‘in time’, i.e. within a second. On the screen, the user is provided with suitable wording for such a turn, e.g. ‘That’s a good idea, let’s talk about it later.’ The user receives feedback on whether or not they responded in time.

In Phase 3 of Module 3, there is practice in starting a turn before the previous speaker has finished, in order to interrupt. The instruction to the user is as follows: ‘When you click the “play” button, the first speaker will speak. You are the second speaker. Take your speaking turn *before* you think the first speaker has finished, to try to steal the floor.’ The user is supplied with an on-screen written prompt for their turn, e.g. ‘That’s a good idea – let’s talk about it later’. A user who successfully completes this phase should be able to take a turn before the current speaker has projected the end of their own turn (i.e. before a turn transition relevance place), in overlap, in order to make a bid for the floor.

Finally, in Phase 3 of Module 4, the user can practise dropping out when accidentally in overlap with another participant. When the user clicks the ‘play’ button, the first speaker will speak. The user takes the role of a potential next speaker. If no-one else speaks immediately, the user is expected to take a speaking turn. For this turn, the user reads a written prompt on the screen e.g. ‘That’s a good idea – let’s talk about it later’. Sometimes the user will find herself in overlap with another speaker. If so, the user is required to drop out immediately. If there is no overlap, however, the user should carry on speaking to the end of the prompt.

## Discussion

This paper has described the development of a self-administered software programme that focusses on participation in everyday conversation, and specifically in handling the problems raised when two participants speak at the same time. The research team was multidisciplinary, representing four different clinical and academic disciplines. The target client group is CI users and for that reason a small group of implant users was closely involved not merely at the start and/or end of the project, as often happens, but throughout the development of the software, as well as its initial dissemination. This approach has meant that consideration of how research findings might be integrated into everyday health service practice is embedded within the project structure. The framework may provide a useful model for involving users, practitioners and researchers together in software development related to conversation skills development.

A further innovative feature of the project was the exclusive use of recorded extracts of naturally occurring conversational speech in the training programme.

This was in order to minimize the gap between practising skills in the training programme and the experience of dealing with overlapping talk in ‘real life’. While this rationale remains valid and after some initial resistance, panel members found that they were able to work with such material, there are some limitations to the work reported here that need to be addressed in future research. One limitation is the focus in the present version of *Talking in Time* on awareness and listening, to the exclusion of speaking activities. The reasons for this have already been discussed. A second limitation relates to the treatment of non-verbal aspects of conversation. While the identification of non-verbal features of overlap types, including gaze direction, gesture and bodily posture, may prove particularly valuable for conversational participants who have a hearing impairment, this dimension was not targeted specifically when developing the *Talking in Time* programme. This was in part because of the paucity of basic research on non-verbal aspects of overlap and in part because of limitations of the video material available to the research team. The recordings had been made originally for a project that focussed on the auditory aspects of overlap and prioritized high-quality single channel audio recordings for each speaker. The purpose of the video recordings was primarily to enable speaker identification. When developing communication-focussed training software in future, it will be important to give equal consideration to both video and audio recordings in order to make full use of the cues used by the participants to manage their participation in conversation. Finally, while informal feedback suggests that *Talking in Time* can be helpful for cochlear implant users, as yet there has been no formal evaluation of its efficacy in enhancing participation in conversation.

## Conclusion

The project was characterized by two innovative features: the involvement of service users throughout the course of the project, not only at its outset or conclusion; and the exclusive use of naturally occurring conversational speech in the training programme. While both present substantial practical challenges, this project has shown that their potential for ecologically valid speech rehabilitation training is considerable. Although the only potential users involved in the development of *Talking in Time* were adults using CIs, there is no reason in principle why it should not prove helpful for people using hearing aids and indeed for people with other types of communication difficulties that impact on their ability to participate fully in conversational interaction. Evaluation of the usefulness of *Talking in Time* for the various types of potential user would be a valuable next step.

## Acknowledgements

The authors would like to thank the following: the five members of the CI user panel for their enthusiastic participation in this project; Dr Suzanne Beeke (UCL) and Dr Caroline Haw (University of Sheffield) for sharing their expertise; Elena Moore and Ella Page for their assistance.

## Disclaimer statements

**Contributors** None.

**Funding** This project was funded by grant AH/L009307/1 from the Arts and Humanities Research Council of the United Kingdom through the scheme 'Follow on Funding for Impact and Engagement'.

**Conflicts of interest** None.

**Ethics approval** None.

## Supplemental material

Supplemental data for this article can be accessed here <https://doi.org/10.1080/14670100.2019.1625185>.

## Notes on contributor

**Bill Wells** is emeritus professor at the University of Sheffield and honorary fellow of the Royal College of Speech and Language Therapists. He has a BA (Hons) in Modern Languages from the University of Oxford and a DPhil in Linguistics from the University of York. After lecturing in the School of Speech Therapy at Birmingham Polytechnic he became Principal of the National Hospitals College of Speech Sciences and subsequently Professor in Clinical Linguistics at University College London. He has published extensively in the field of interactional phonetics and phonology as well as on children's typical and atypical speech and language development. With Joy Stackhouse he developed a psycholinguistic framework for assessing children with speech and literacy difficulties.

**Amy Beeston** is a Visiting Academic in Computer Science at the University of Sheffield, and a Visiting Fellow in Music at the University of Leeds, UK. She holds a PhD in Computer Science (University of Sheffield), a Masters in Sonology (The Royal Conservatoire, The Hague) and a BMus (Hons) in Music Technology (Edinburgh). Her research focus is on human and machine listening, and she works across scientific and artistic communities with partners in educational, clinical and industrial settings.

**Erica Bradley** was the Specialist Speech and Language Therapist for the Adult Cochlear Implant Assessment and Rehabilitation Service at Sheffield Teaching Hospitals NHS Foundation Trust from 2008 to her retirement in 2018. Erica undertook her undergraduate degree at the University of Sheffield in 1994 and was a member of the Royal College of Speech

and Language Therapists. Erica has extensive experience of working with individuals with highly complex communication difficulties, alongside wide experience of assessment and management of dysphagia in acute, rehabilitation, and community settings. From 2004 she held the role of Speech & Language Therapy Coordinator for the Adult Service, Sheffield PCT. Erica has also been a Health Professions Council Partner for Sheffield Teaching Hospitals NHS Foundation Trust and a member of NICE topic expert group developing Stroke Quality Standards.

**Guy J. Brown** obtained a BSc (Hons) Applied Science from Sheffield City Polytechnic in 1984 and a PhD in Computer Science from the University of Sheffield in 1992. He was appointed to a Chair in the Department of Computer Science, University of Sheffield, in 2013. He has held visiting appointments at LIMSI-CNRS (France), Ohio State University (USA), Helsinki University of Technology (Finland) and ATR (Japan). His research interests include computational auditory scene analysis, speech perception, hearing impairment and acoustic monitoring for medical applications. He has authored more than 100 papers and is the co-editor (with Prof. DeLiang Wang) of the IEEE book "Computational auditory scene analysis: principles, algorithms and applications".

**Harriet Crook** is The Organisational Lead for Healthcare Science at Sheffield Teaching Hospitals NHS Trust. She has worked as a Clinical Scientist in the NHS since 2003 across Adult Rehabilitation, Auditory Processing Disorders and Cochlear Implants. She completed a PhD in Auditory Neuropsychology (University of Sheffield, 2000) and an M.Sc. in Audiological Science (University of Manchester, 2001). She is currently Chief Examiner for the British Academy of Audiology. Dr Crook is actively engaged in funded research and holds collaborations with the Departments of Music, Computer Science and Human Communication Sciences at the University of Sheffield and with the School of Music at the University of Leeds. She has been clinical lead for the recent AHRC funded 'Hearing Aids for Music' Project and has been part of the Sheffield Teaching Hospitals 'In and out of Hospitals' work, funded by the Arts Council, providing music workshops for those with hearing loss. Her research explores Auditory streaming, Auditory processing in Cochlear Implant users and music listening with implants and hearing aids. She has a personal interest in music and hearing loss via her undergraduate degree in Music and Philosophy and becoming a cochlear implant recipient herself in 2018.

**Emina Kurtić** is an Honorary Research Fellow at the Department of Human Communication Sciences at the University of Sheffield. She holds a PhD from the University of Sheffield, an MSc in Speech and



Language Technologies from the University of Edinburgh and an MA in Linguistics, Mathematics and Economics from Ruhr-University Bochum (Germany). Her research is focused on naturally occurring human conversations, and she works with large amounts of conversational data from a range of domains including research meetings, everyday conversations, online commenting forums and conversations in clinical setting. Her research uses a combination of Conversation Analysis and Data Mining to discover and describe conversational phenomena in various languages.

## ORCID

Amy V. Beeston  <http://orcid.org/0000-0003-2796-1947>

Guy J. Brown  <http://orcid.org/0000-0001-8565-5476>

## References

- Amann, E., Anderson, I. 2014. Development and validation of a questionnaire for hearing implant users to self-assess their auditory abilities in everyday communication situations: the hearing implant sound quality. *Acta Oto-Laryngologica*, 134: 915–923.
- Anstey, J., Wells, B. 2013. The uses of overlap: carer-child interaction involving a nine-year-old boy with auditory neuropathy. *Clinical Linguistics & Phonetics*, 27(10–11): 746–769.
- Auer, P. 2017. *Gaze, addressee selection and turn-taking in three-party interaction* (INLIST – Interaction and Linguistic Structure No. 60). Retrieved from <http://www.inlist.uni-bayreuth.de/issues/60/index.htm>.
- Beeke, S., Sirman, N., Beckley, F., Maxim, J., Edwards, S., Swinburn, K. et al. 2013. Better conversations with aphasia: an e-learning resource. Retrieved from [www.ucl.ac.uk/betterconversations/aphasia](http://www.ucl.ac.uk/betterconversations/aphasia).
- Crew, J., Galvin, J.J., Fu, Q. 2016. Perception of sung speech in bimodal cochlear implant users. *Trends in Hearing*, 20: 1–15.
- Egbert, M. 1997. Schisming: the collaborative transformation from a single conversation to multiple conversations. *Research on Language and Social Interaction*, 30(1): 1–51.
- French, P., Local, J. 1983. Turn-competitive incomings. *Journal of Pragmatics*, 7: 17–38.
- Fuller, C.D., Gaudrain, E., Clarke, J.N., Galvin, J.J., Fu, Q.J., Free, R.H. et al. 2014. Gender categorization is abnormal in cochlear implant users. *JARO – Journal of the Association for Research in Otolaryngology*, 15: 1037–1048.
- Heldner, M., Edlund, J. 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, 38(4): 555–568.
- Hilton, K. 2016. The perception of overlapping speech: effects of speaker prosody and listener Attitudes. In: *Interspeech 2016*. San Francisco: ISCA, p. 1260–1264. doi:10.21437/Interspeech.2016-1456.
- Jefferson, G. 1984. Notes on some orderlinesses of overlap onset. In: D'Urso V., Leonardi P., (eds.) *Discourse analysis and natural rhetoric*. Padua: Cleup Editore, p. 11–38.
- Jefferson, G. 1987. Notes on “latency” in overlap onset. *Human Studies*, 9: 153–183.
- Jefferson, G. 1989. Preliminary notes on a possible metric which provides for a “standard maximum” silence of approximately one second in conversation. In: Roger D., Bull P., (eds.) *Conversation: an interdisciplinary perspective*. Clevedon: Multilingual Matters, p. 166–196.
- Kovačić, D., Balaban, E. 2009. Voice gender perception by cochlear implantees limit cycle oscillations in a nonlinear state space model of the human cochlea. *The Journal of the Acoustical Society of America*, 126(118): 2124–2751.
- Kurtić, E. 2011. *Overlapping talk and turn competition in multi-party conversations* [PhD Thesis]. University of Sheffield.
- Kurtić, E., Brown, G.J., Wells, B. 2010. Resources for turn competition in overlap in multi-party conversations: speech rate, pausing and duration. In: *Interspeech-2010*, p. 2550–2553.
- Kurtić, E., Brown, G.J., Wells, B. 2013. Resources for turn competition in overlapping talk. *Speech Communication*, 55(5): 721–743.
- Kurtić, E., Gorisch, J. 2018. F0 Accommodation and turn competition in overlapping talk. *Journal of Phonetics*, 71: 376–394.
- Kurtić, E., Wells, B., Brown, G.J., Kempton, T., Aker, A. 2012. A corpus of spontaneous multi-party conversation in Bosnian Serbo-Croatian and British English. In: Calzolari N., (ed.) *Proceedings of the 8th international conference on language resources and evaluation, LREC 2012*. Istanbul: European Language Resources Association (ELRA).
- Lee, C., Lee, S., Narayanan, S.S. 2008. An analysis of multimodal cues of interruption in dyadic spoken interactions. In *INTERSPEECH 2008 – 9th Annual Conference of the International Speech Communication Association*, Brisbane, Australia, pp. 1678–1681.
- Levinson, S.C., Torreira, F. 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6: 1–17. (Article 731).
- Litovsky, R.Y., Parkinson, A., Arcaroli, J. 2009. Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear and Hearing*, 30(4): 419–431.
- Luo, X., Masterson, M.E., Wu, C.-C. 2014. Contour identification with pitch and loudness cues using cochlear implants. *The Journal of the Acoustical Society of America*, 135(1): EL8–EL14.
- Max. 2012. Walnut, CA: Cycling '74.
- Mondada, L., Oloff, F. 2011. Gestures in overlap: the situated establishment of speakership. In: Stam G., Ishino M., (eds.) *Integrating gestures: the interdisciplinary nature of gesture*. Amsterdam: John Benjamins Publishing Company, p. 321–338.
- Most, T., Peled, M. 2007. Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. *Journal of Deaf Studies and Deaf Education*, 12 (3): 350–361.
- O'Halpin, R. 2010. *The perception and production of stress and intonation by children with cochlear implants* [PhD Thesis]. University of London. Retrieved from <http://ethos.bl.uk>.
- Peirce, J. 2014. PsychoPy. Retrieved from <http://www.psychopy.org>.
- Peng, S.C., Lu, N., Chatterjee, M. 2009. Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiology and Neurotology*, 14(5): 327–337.
- Peng, S., Tomblin, J.B., Turner, C.W. 2008. Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing*, 29(3): 336–351.
- Pomerantz, A., Heritage, J. 2012. Preference. In: Sidnell J., Stivers T., (eds.) *Handbook of conversation analysis*. Chichester: Wiley-Blackwell, p. 210–228.
- PowerPoint. 2011. Redmond, WA: Microsoft.
- Sacks, H., Schegloff, E., Jefferson, G. 1974. A simplest systematics for the organisation of turn-taking. *Language*, 50(4): 696–735.
- Schegloff, E. 2000. Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, 29(1): 1–63.
- Schegloff, E.A. 2001. Accounts of conduct in interaction. In: J.H. Turner, (ed.) *Handbook of sociological theory*. New York: Kluwer Academic/Plenum Publishers, p. 287–321.
- Skelt, L. 2013. Recipient gaze and the resolution of overlapping talk in hearing impaired interaction. *Journal of Interactional Research in Communication Disorders*, 4(1): 71–94.
- Snow, D., Ertmer, D. 2012. Children's development of intonation during the first year of cochlear implant experience. *Clinical Linguistics & Phonetics*, 26(1): 51–70.
- Tannen, D. 1983. When is overlap not an interruption? One component of conversational style. In: Di Pietro R., Frawley W., Wedel A., (eds.) *Selected Papers of the First Delaware Symposium on Language Studies*. Newark: University of Delaware Press, p. 119–129.
- Tye-Murray, N., Witt, S. 1996. Conversational moves and conversational styles of adult cochlear-implant users. *Journal of the Academy of Rehabilitative Audiology*, 29: 11–26.
- Van De Velde, D.J., Dritsakis, G., Frijns, J.H.M., Van Heuven, V.J., Schiller, N.O. (2015). The effect of spectral smearing on the identification of pure F0 intonation contours in vocoder simulations of cochlear implants. *Cochlear Implants International*, 16(2): 77–87.
- Varley, R., Cowell, P.E., Dyson, L., Inglis, L., Roper, A., Whiteside, S.P. 2016. Self-administered computer therapy for apraxia of speech: two-period randomized control trial with crossover. *Stroke*, 47: 822–828.
- Wittenburg, P., Brugman, H., Russel, A., Klassman, A., Sloetjes, H. 2006. ELAN: a professional framework for multimodality research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.