University of Nottingham
UK | CHINA | MALAYSIA

Operationalising success in group conversations for people with normal and impaired hearing

Raluca Nicoras

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Hearing Sciences - Scottish Section
Division of Clinical Neuroscience, School of Medicine
University of Nottingham
2024

I declare that this thesis has been composed by myself and that the research reported here has been conducted by myself unless otherwise indicated. This work has not been submitted for any other degree or professional qualification.

Raluca Nicoras
Glasgow, May 17th, 2024

Abstract

Everybody wants to experience success in their conversations. However, people with hearing loss face challenges, especially in group settings. While hearing aids promise enhanced listening experiences, little is known about their impact on real-life conversation situations. There is a pressing need to incorporate conversation success into measures of hearing aid outcomes. The lack of a clear definition of conversation success complicates our attempts to understand hearing aid benefits in conversations. Therefore, the research presented in this thesis investigates what constitutes conversation success for individuals with normal and impaired hearing, focusing specifically on group conversations.

I start by providing an overview of the literature on conversation, focusing on aspects relevant to conversation success. This includes examining theoretical models of human communication (such as linear, interactive, and constitutive models), specific mechanisms that may relate to conversation success (such as behaviour rules, entrainment, and accommodation), characteristics of the interacting group that could influence conversation success (such as group size, conversation goal, and familiarity among interlocutors) and methodological approaches in conversation research. Then, I present findings from a series of three studies aimed at understanding and operationalising conversation success for individuals with both normal and impaired hearing from three different perspectives:

Subjective perspective (Study 1): *What is conversation success for people with normal and impaired hearing?*

Older adults with normal and impaired hearing wrote down ideas about conversation success. They sorted and then rated these ideas in terms of their importance in one-to-one and group conversations. Results showed that conversation success is given by "*Being able to listen easily", "Being spoken to in a helpful way", "Being engaged and accepted", "Sharing information as desired", "Perceiving flowing and balanced interaction", "Feeling positive emotions*" and "*Not having to engage coping mechanisms*". Respondents considered that "*Being able to listen easily", "Being spoken to in a helpful*

way" and "*Sharing information as desired"* are more important in group than in one-to-one conversations.

Behavioural perspective (Study 2): *What behaviours are associated with successful conversations?*

A survey conceived based on the seven themes resulting from the previous study was used to evaluate the perception of conversation success in a face-to-face conversation experiment involving 18 groups of 4 people (quartets). Each quartet was composed of two people with normal hearing and two people with impaired hearing (hearing aid users) who held six conversations across low, medium, and high levels of background noise. Participants with impaired hearing were unaided in half of the conversations. Participants' vocal activity, head movement and perception of conversation success were recorded during each conversation. Linear mixed models were used to link participants' perceptions of conversation success to their behaviours. Findings showed that both vocal activity and head movements are associated with the perception of conversation success.

Third-party observer perspective (Study 3): *How do third-party observers judge conversation success?*

Naive observers watched eight two-minute snippets of prerecorded conversations. Their task was to assess the success of these conversations when presented as visual-only versus auditory-visual videos and to provide the cues that informed their assessment. Most responses from the observers reflected the seven original success themes, with all original themes except one (*Not having to engage coping mechanisms*) represented in the observers' responses. Additionally, two new themes emerged: observers considered a conversation successful when participants were *Engaging effort positively* and *Bonding over difficulty*. In terms of ratings, when observers viewed visual-only videos of unsuccessful conversations, they rated these conversations as more successful than they actually were.

These findings provide different perspectives on conversation success, combining self-perception, communication behaviours and other-perceptions. Taken together, these studies suggest that there are nine factors related to

conversation success, that vocal activity and head movement are associated with conversation success, and that auditory input is critical in determining conversation success, particularly in identifying unsuccessful conversations.

Acknowledgements

This research was supported by WSAudiology.

First, I would like to thank my (many!) supervisors: Graham, Lauren, Karolina, Rosa-Linde, and Sarah. This thesis would not have been possible without your guidance and support. Thank you for your time, patience, and also for never laughing at my overly ambitious and unrealistic deadlines. I have learned so much in our countless one-to-one and never-ending 'all supervisors' meetings. Your input shaped my research and led to this thesis.

Graham, I am incredibly grateful for your invaluable guidance and mentorship. You have always made yourself available to answer my questions whenever I knocked on your door (often without scheduling a meeting!). Not only did you support me in this project, but you also sent me across the ocean to pursue my research interests (Canada), transformed my vague ideas into real grant proposals, and encouraged my creativity in side projects (HSSS Christmas video).

Lauren, your expertise in conversation behaviours has been incredibly valuable. Your passion for (quantitative!) research and your quick wit inspired me throughout my PhD journey. With keen insights and clever humour, you made our conversations so enjoyable. Thank you for your emotional support. I appreciated your ability to recognise when I wasn't feeling okay, even when I didn't realise it myself. Also, I learnt so much about academia from you. Thank you for everything.

Karolina, you've been my sparring partner in many areas! In my first year, you guided me through what I needed to know about hearing loss and hearing aids. You also gave me the opportunity to co-supervise two BMedSci students and supported me in the (very long) qualitative analysis of Study 3. Thank you for making our collaboration productive and enjoyable. I look forward to continuing our tradition of watching operas together!

Rosa-Linde, thank you for always being straightforward in our all-supervisors meetings and for keeping the group on track. Sarah, I truly appreciate your openness to new methodological approaches, and I am grateful for your input in Study 1.

I am grateful to the ORCA team for hosting me in their lab in Stockholm, to Jenny Campos and her team at KiTE for their hospitality, to Haim Bar for his guidance in GCM methods, to Oliver Zobay for his assistance, and to Tim Beechey for teaching me how to review papers and for his support in OPA.

I would also like to thank all my colleagues at HSSS who made my workplace such a friendly environment. I want to mention Andrew Lavens, who offered technical support in all studies of my PhD, and all HSSS PhD students (Janin, Eszter, Diana) who never judged the inhuman amount of chocolate I consumed during these years.

Special thanks to Matthew, who has offered the most support both in and outside of work. Meeting him has been one of the best outcomes of this PhD.

Outside of the department, I thank my friends from Romania for embracing our long-distance friendship and always welcoming me back with open arms. Also, I want to thank Victor, who put a roof over my head when I needed one, patiently listened to me rehearse my first conference talk at midnight on Skype and spent Saturday afternoons explaining regressions to me.

I want to thank my extended family (Luci, Lili, Antonia, Vlad, Cristi, and Adelina), as well as my brother and sister-in-law (Andrei and Adriana), who have always supported me and kept me grounded in reality during these years. They reminded me that there is life outside of work. A special thank you to my niece, Emma, whose inspiring curiosity and contagious laughter have brought so much joy into my life.

Finally, I would like to thank my parents, Vasile and Mihaela, for their unconditional support, love, and affection. Throughout these years, they cheered for every talk, poster presentation, prize, and publication as if it were the most significant achievement. Their happiness means the world to me.

Table of Contents

List of tables

List of Figures

Chapter 1
Introduction

1.1 Chapter summary

Why investigate conversation success in hearing sciences? There are two primary reasons that this topic is of interest to both the academic world and the hearing aid industry. Firstly, everybody aspires to have successful conversations, yet individuals with hearing loss often encounter challenges. While hearing aids provide improvements in sound quality and enhanced listening experiences, their efficacy in real-life communication scenarios remains limited. Therefore, it is essential to better understand how hearing aids perform in realistic communication contexts and how they contribute to conversation success. Such insights could inform the design of hearing aids, ultimately improving the quality of life for individuals with hearing loss. Secondly, understanding conversation success opens new avenues for developing helpful communication strategies and training frequent conversational partners to improve their communication techniques. This facilitates successful interactions with individuals experiencing hearing loss, improving overall conversation experience for all parties involved.

In this chapter I provide an overview of the literature on conversation specifically focusing on aspects relevant to conversation success. Initially, I look at how human communication evolved into its present form (section 1.3.1). Then I explore the processes involved conversation as outlined in various theoretical models of human communication under the premise that when all processes involved in a conversation function optimally, it has the potential to lead to conversation success (section 1.3.2). While some theoretical models acknowledge that in conversation the information flows bidirectionally rather than solely in a linear fashion (from transmitter to receiver), they focus on dyadic interactions rather than considering larger group dynamics. Next, I explore several characteristics that could impact conversation success, beginning with the size of the group (1.3.3) to provide context for my focus on group conversations in Chapters 3, 4, 5, and 6. I also look at various

methodologies employed in measuring conversation, ranging from field experiments to laboratory studies and ecological momentary assessments (1.3.4), as well as providing a review of the processes that facilitate human communication as background for my results in Chapters 4, 5, and 6. Then, I look at the notion of conversation success and the criteria that could determine such a judgment. The second part of this chapter focuses on the population targeted in my research: people with hearing loss. Here, I offer an overview of how existing evidence concerning hearing loss affects conversation and describe the role played by hearing aids in mitigating these challenges. The chapter concludes by summarizing the research aims of this dissertation and presenting a chapter plan.

1.2 Motivation

This thesis explores Conversation Success as perceived by individuals with normal and impaired hearing. The central motivation stems from the recognition that conversation, deeply embedded in our evolutionary behaviour, constitutes a significant aspect of our daily lives. Despite its apparent simplicity, engaging in conversation involves simultaneous use of multiple processes such as listening, comprehending, preparing a response, and predictions. The complexity of this task makes it susceptible to various challenges, among which hearing loss is a very prevalent one. Traditionally, the processes involved in conversation are individually scrutinized; neuroscience focuses on the brain, and hearing studies centre on the ear and listening processes. However, in our everyday experiences, hearing serves not only for listening but as an integral component of our communication system. Individuals with hearing loss often seek audiological assistance due to difficulties in understanding speech, emphasizing the role of hearing in conversations.

Currently, hearing assessments focus on listening tests. However, this thesis argues for a shift towards examining hearing from a conversational perspective. Rather than exclusively relying on traditional listening tasks, which focus on auditory perception in controlled settings, the studies conducted in this thesis immerse participants in conversational contexts.

Additionally, existing research typically focuses on challenging conversational situations, where reactions to obstacles provide valuable insights. While acknowledging the potential lessons from overcoming obstacles, the effects of hearing loss extend beyond coping mechanisms. Individuals seek not only to cope with their hearing loss but to restore their abilities to a pre-loss state. Therefore, the ability to hold successful conversations could be one measure of the quality of their hearing rehabilitation.

1.3 Background - conversation

The etymology of the word "conversation" is rooted in Latin, originating from "conversatio," a fusion of "con-" (with) and "versari" (to turn or to be involved in), meaning "to live with" or "to keep company with". Already from its origins, the nature of the term suggests that it is not a static label like a word that describes an object (e.g.: “book”), but rather a term signalling an action or a range of actions. This idea is not an original one, instead, it is inspired by a popular opinion in the field of Organizational Behaviour and Psychology. Karl Weick (1979) advocated for the use of the term “organising” instead of “organisation” arguing that the latter is too static to capture the dynamic processes at play. Similarly, conversation is not a given, it does not simply exist without a multitude of actions that constitute it. Although “conversing” might be better suited, I will consistently use the term "conversation" as an umbrella for all the actions that make conversation possible. In the following sections, I will explore these actions, first from an evolutionary perspective, then through the theoretical models of human communication, and then through the mechanisms that facilitate conversations.

1.3.1 The evolution of face-to-face communication

Speaking, listening, and understanding are only a few actions that are known to be part of the “interaction engine” (Levinson et al. 2006). A very early version of this engine predates the form of language as we know it today and emerged over half a million years ago, likely driven by natural development and potentially influenced by ecological changes that required increased cooperation (Levinson, 2020). Evolutionarily, the transition from

basic gesture-based language to complex vocal communication mirrors the developmental stages observed in children's communication. They both start from a very simple base (e.g.: gesture-based language) with subsequent layers added without replacing previous ones (e.g.: spoken language acquisition complements rather than replaces hand gestures). The sequential accumulation of communication layers contributes to the development of a multi-modal communication system that integrates gesture and speech (Levinson et al., 2014). Our communicative skills are also supported by the evolution of physical characteristics: humans have highly expressive faces which, compared to other primates, enable visibility of facial expressions; dexterity due to the evolution of the opposable thumb and fine motor control; stereoscopic vision that provides depth perception aiding communication through a better interpretation of visual cues; the descent of the larynx which, although debated (Fitch & Reby, 2001), seems to have appeared as a facilitator of articulate speech; human ear selectivity for speech frequencies (Manley, 2017), and upright posture facilitating face-to-face interaction. The ability to combine our physical characteristics to convey multi-modal messages confirms that human communication is a system of systems (Levinson et al., 2014). Conversation can take various forms including political debates, written messages and virtual conversations, the focus of this thesis remains on face-to-face in person verbal conversations. In a period increasingly dominated by digital communication, face-to-face in person conversation remains a reference point for multiple fields: psychology, human-computer interaction, management, and hearing sciences. The physical proximity between participants facilitates multimodal communication through the natural integration of informative auditory and visual cues such as head movements (Ishi et al., 2014) gestures (Wagner et al., 2014), facial expressions (Frith, 2009), and lip reading (Sumby & Pollack, 1954). Evolutionarily speaking, we are fully equipped for successful conversations. But possessing the necessary tools does not guarantee success.

1.3.2 Theoretical models of communication

The study of communication processes occurs at both theoretical and practical levels. Theoretical models play a role in shaping communication research, and conversely, research outcomes provide supporting evidence for these models or inform the development of new ones. These models use symbolic representations to break down the complexity of the communication process, simplifying its components for better understanding. Despite their usefulness in providing general guidance, models may nonetheless overlook essential aspects of communication. The three main types of theoretical models of communication are: linear, transactional/interactional, and constitutive. It is worth noting that communication and conversation are related concepts, but have distinct meanings. While conversations are a form of communication, communication can encompass a broader range of interactions (i.e.: public speaking). The purpose of this section is to portray examples of theoretical models of human communication and explore their evolution through the lenses of conversation success.

1.3.2.1 Linear models of communication

Linear models, as suggested by their name, represent a unidirectional process, where a message is sent from a sender to a receiver without considering the presence of a response. While there might be multiple linear models, we describe below the most popular ones:

Shannon-Weaver Model (1948) (Figure 1.1): Developed by Claude Shannon and Warren Weaver while working for Bell Telephone Laboratories, this is one of the most popular theoretical models. It breaks down the communication process into specific elements that follow a certain sequence (Table 1.1).

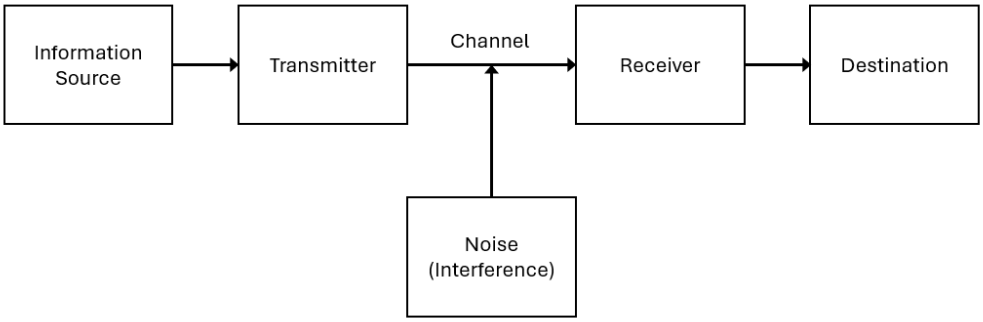


Figure 1.1 A schematic diagram illustrating Shannon-Weaver (1948) model.

Table 1.1 Shannon-Weaver (1948) model elements and their definitions

Element	**Definition**
Information source	The person or entity originating the message.
Transmitter	The process of converting the sender's ideas into a message.
Message	The information or content being transmitted.
Channel	The medium or pathway through which the message is sent.
Noise (inference)	Interference or disruptions that can affect the message during transmission.
Receiver	The process of interpreting the received message.
Destination	The person or entity for whom the message is intended.

The model was widely used in technical communication, considering transmission through machines. However, it has various limitations which include omission of other variables that affect the communication process such as feedback and limited applicability to interpersonal and group conversation.

Harold D. Lasswell (1948) presents a model with a very similar structure but that is primarily designed for mass communication. While this form of communication is out of the scope of this thesis, it is noteworthy to highlight that Laswell's model does not conclude with the "receiver", but extends to the effect that the communication has on the "receiver" (table 1.2).

Table 1.2 Lasswell's Model elements and definitions

Element	Definition
Who	Identifies the sender or source initiating the communication.
Says What	Describes the content or message being communicated.
In Which Channel	Specifies the medium or pathway used for transmitting the message.
To Whom	Identifies the audience or receiver for whom the message is intended.
With What Effect	Examines the impact or consequences of the communication on the receiver.

The "Source-Message-Channel-Receiver" model (Berlo, 1960) also operates on a linear trajectory, with communication flowing in one direction from the Sender to the Receiver. However, this model takes into account that the transmission process can be influenced by the source and the receiver's characteristics. It explores factors affecting communication efficiency, including communication skills, attitudes, knowledge, social systems, and culture (Figure 1.2). It also recognises that the message can be influenced by various factors such as how, when, and where it is transmitted. Similarly, the transmission channel can be multilayered; participants can send and receive messages through all sensory cues, including sight, hearing, touch, smell, and taste. While Berlo's model is closer to be useful for face-to-face conversations, it does not account for reciprocity.

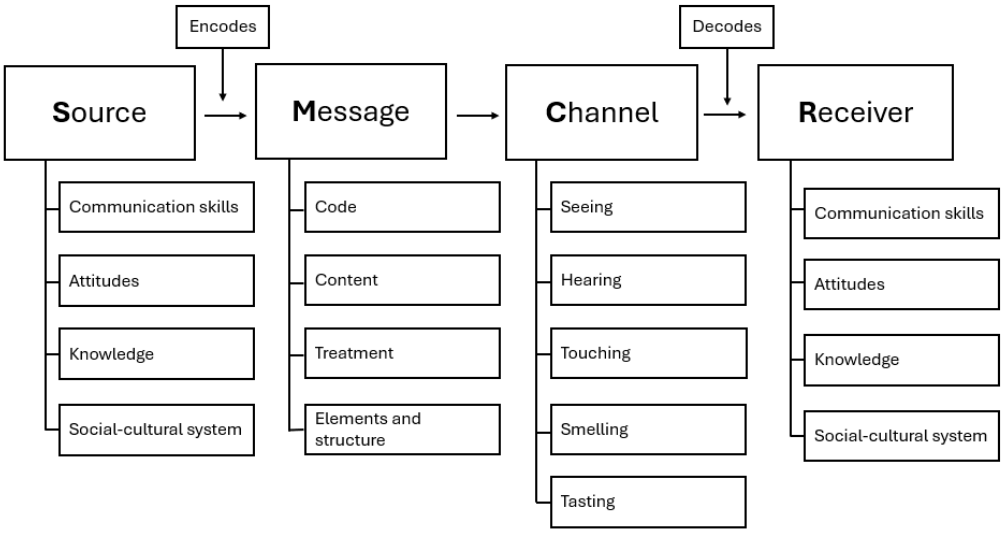


Figure 1.2 Schematic diagram showing Berlo's model of communication.

Although not designed for this purpose, these communication models could be interpreted to have a very narrow view of what constitutes successful communication. According to the linear models, a successful face-to-face communication process takes place when Person A has a message to send, encodes this message through one of the behavioural modalities available (e.g., speech, gesture, facial expression), the content of the message is passed to person B while overcoming the potential interferences (e.g., noise), and then Person B accurately decodes the message. Yet, when attempting to apply this model to real-life scenarios, it becomes evident that the complexities of conversation surpass the simplified representations provided by the model. Misunderstandings and incomplete information transfer are common in natural conversation, yet communication is largely remarkably successful. Indeed interlocutors are able to use a variety of techniques to succeed. For example, a listener in a conversation doesn't listen motionless and silent until a talker has finished. Conversation is a two way or multi-way process, a dialogue rather than a monologue (Brennan, 2016). The exchange of information is not strictly unidirectional; both senders and receivers have the capacity to transmit information simultaneously and can also interchange their roles. Linear models seem to minimise the sender role while maximising the receiver competences. It also assumes that sender can and has the intention to convey the message to the best of his abilities so that it can be received and understood by the receiver. It also assumes both sender and receiver have a common ground (Stalnaker, 2002), which can be translated to them speaking the same language, or simply having common knowledge. The sender decides consciously (or not) how to encode the message, by making use of specific behavioural modalities. The informational content is not only in the content of the message but also in the sensory information displayed by the sender. It is argued that conversing humans are not only sending a message, but they are also expressing their attitude about the content of the message (Frick, 1985), which can be inferred through tone, gestures, or facial expressions. Thus while these models consider success to lie entirely on information transfer and require interlocutors to plan

and convey perfect versions of their messages, conversation success might be more than a linear transmission of information in which all components work perfectly.

1.3.2.2 Interactive/transactional models of communication

While some researchers treat interactive models and transactional models as different, I chose to describe a representative model that synthesizes key elements from both, considering them similar enough to be presented together for the purpose of this thesis. The interactive models consider communication as a two-way, dynamic process involving simultaneous sending and receiving of messages. They incorporate feedback and therefore demonstrate greater relation to face-to-face conversation compared to linear models. Schramm (1954) introduced the idea of a feedback loop, pointing out that a participant can be at the same time: source, encoder, interpreter, and decoder (figure 1.3).

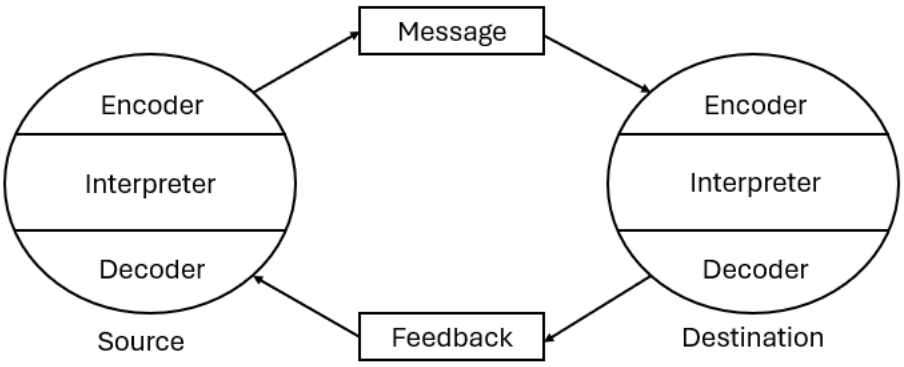


Figure 1.3 A schematic diagram showing a representative Interactive/transactional model as described by Schramm (1954).

The model implies that communication can be bidirectional; however, it falls short in addressing additional dimensions evident in real-life communication settings. For instance, the presumption that communication follows a strict cyclical pattern may not hold true, as in actual conversations, individuals often talk at the same time. Also, this model overlooks essential

factors such as non-verbal messages and noise (barriers). All these might play pivotal roles in shaping successful communication.

1.3.2.3 Constitutive model of communication

While earlier models primarily view communication as a tool for conveying information such as knowledge, social meaning and emotions, Craig (1999) introduces the idea that communication is the driving force behind our social lives rather than an exchange of information. As articulated by Manning (2014): “communication is not a mere tool for expressing social reality, but also a means of creating it”. This communication (meta)model transcends the reduction of the process into its constituent elements, focusing instead on its role in shaping our social reality. This perspective is parallelled by studies showing how conversation facilitates relationship formation and reputation management (Dunbar et al., 1997; Emler, 1990). Therefore, conversation success might not be only related to the optimal functioning of the communication components, but also to its functionality in individuals’ social lives.

1.3.2.4 Summary - theoretical models of communication

Theoretical communication models present varied perspectives for approaching the study of communication and can serve as inspiration for research questions and experimental paradigms. Linear models of communication might correspond to basic listening paradigms (i.e.: individuals are asked to passively listen to speech/sound) due to sender–passive receiver conceptualisation. Interactive models of communication might correspond to interactive paradigms that allow listener feedback (e.g.: Diapix task). Finally, the constitutive model of communication might correspond to decision-making paradigms where conversation can be used, for example, as a tool of persuasion shaping individuals’ social lives.

However, these models do not offer a complete picture of conversation success. For instance, they overlook individuals' desires or ability to engage and sustain the communication process. Additionally, they lack elements related to individuals' expectations, goals, or desire for successful

communication. Finally, while they seem to illustrate the components of dyadic conversations, they do not easily extend to the complexities of conversations with more than two participants. Next, I examine several characteristics of conversation that could impact participants' perception of conversational success.

1.3.3 Conversation characteristics

From one-to-one intimate conversations between partners to business meetings between collaborators, conversations can take various forms. Below, I explore several key factors that define a specific conversation type or situation. I also try to provide an overview of their prevalence in our daily lives.

1.3.3.1 Size

Conversation success might depend on the size of the group. In dyadic conversation, the responsibility for success is divided into two, whereas in group conversations, the responsibility for success is distributed among multiple participants. While this might alleviate the individual burden to make the conversation successful, it might also mean that individuals have less control over the conversation's direction and outcome.

Previous research has largely focused on conversations between two people (dyads), although dyads may only account for about 50% of all conversations (Dunbar et al., 1995; Peperkoorn et al., 2020) Group conversations, defined as conversations among more than two people, are part of our daily lives as well.

In terms of communication behaviour, interactions involving more than two people introduce new complexities. While in dyadic interactions each person’s attention is directed towards the only dyadic partner, in group conversations attention needs to be divided between participants. This might have a range of consequences such as increased head movements and gaze variability (Hadley et al., 2021), less predictable speaker sequences and less time to speak (Cooney et al., 2020) hence a more intense competition for taking and holding the floor. Groups also offer a larger repertoire of possible

behaviours. In dyads there is always the pressure of being the next speaker, while in groups the obligation to speak is removed. This might prompt individuals to take more time to comprehend what they hear or employ withdrawal strategies. The size of the conversing group may significantly influence what is considered successful conversation, as the conversation flow varies with the number of participants. In larger groups, conversations may take the form of interactive dialogues or serial monologues, with one participant assuming the role of the speaker while others listen (Fay et al., 2000). Hence criteria for success could vary dramatically between dyadic and larger group conversations.

1.3.3.2 Goal

The link between conversation success and conversation goals could lie in how effectively those goals are achieved within the conversation. If the participants perceive that their objectives for the conversation have been met, they may consider it successful regardless of other factors such as the quality of communication or participant engagement. Therefore, understanding the types of conversational goals individuals have might help in understanding what constitutes conversation success.

Research on social interaction shows two categories of conversational goals. Relational conversations, characterized by free and open interactions, play a significant role in the human need to bond through conversation. They usually take place between family and friends. These can take the form of deep and affectionate conversations and seem to be linked to improved health and reduced stress, particularly in close relationships (Floyd et al., 2007). This form of communication is generally reserved for relational intimates. These conversations can be characterised by gossip, or joking, and contribute to cementing social bonds and increasing relational satisfaction (Burns & Pearson, 2011; Dunbar et al., 1995; Reis et al., 2018). In contrast, informational conversations are goal-oriented conversations and usually take place in work-related contexts, being motivated by achieving specific objectives. They appear to be negatively associated with feelings of connection among friends. However, they seem to occur frequently (Reis et al., 2018).

Sometimes, conversations can become complex when participants alter their goals mid-conversation, concurrently pursue multiple goals which may conflict, or even hold conversations without being aware of their goals. The conversational circumplex (Figure 1.4) (Yeomans et al., 2022) addresses these issues by combining informational and relational goals as dimensions in a single framework. On one dimension (x axis) are represented the informational goals (giving and receiving information). On the second dimension (y axis) are represented the relational goals (bonding and building relationships). Yeomans et al. (2022) argue that any conversational goal could be positioned somewhere between these two axes. According to them, in order to achieve success during a conversation, individuals only need to be aware of their goals and to align their actions accordingly.

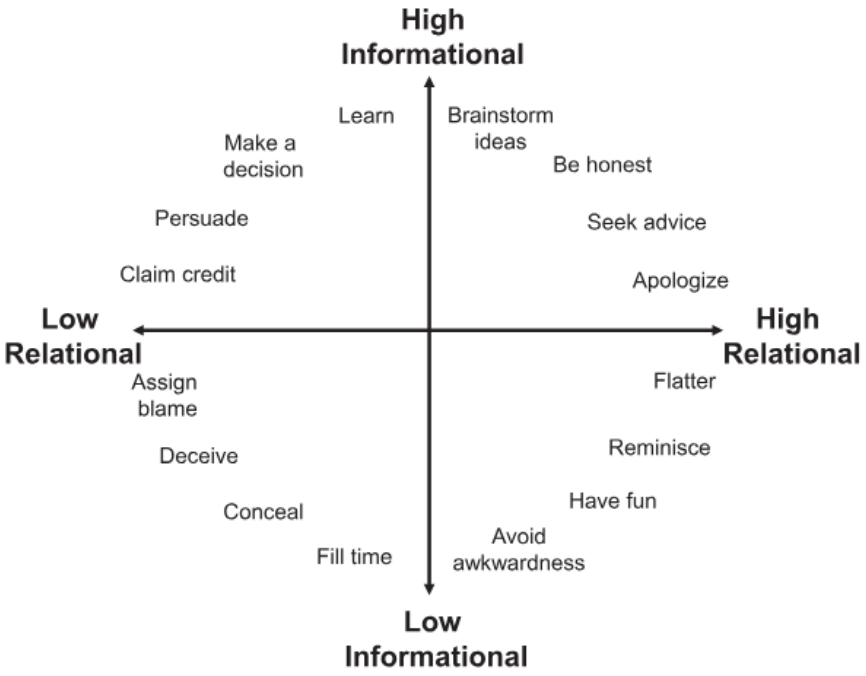


Figure 1.4 The conversational circumplex framework (Yeomans et al. 2022)

1.3.3.3 Familiarity among interlocutors

The perception of conversation success might also differ depending on how familiar the interlocutors are. When communication partners are familiar with each other, they may have established rapport, shared experiences, and a

deeper understanding of each other's communication styles. These factors can lead to different conversational expectations compared to interactions with strangers.

Conversational behaviour can show variations as a function of participants' familiarity. In interactions between individuals who share a close relationship or are well-acquainted, conversations often involve a higher degree of intimacy shown by physical proximity, shared understanding shown by a shorter time needed to achieve grounding, and a more relaxed atmosphere. Participants may employ informal language, inside jokes, and references that are specific to their shared experiences. This corresponds to a higher level of alignment between familiar people. It has been suggested that we are more neurally aligned with friends than with friends of friends, and more aligned with friends of friends than with strangers (Parkinson et al., 2018). Also, research suggests that individuals find it easier to predict the end of a speaker's turn when the speaker's voice is more similar to their own compared to when the speaker is a stranger (Hadley et al., 2020). These findings might translate into a smoother interaction, with fewer repairs, and less overlap. Therefore, communication behaviour that could relate to conversation success might be different depending on interlocutors' familiarity.

All these conversation characteristics (such as size, goal, and familiarity among interlocutors) that could influence how conversation success is perceived are usually taken into account in experiments. While some studies focus solely on one type of conversation to avoid confounding outcomes (e.g., measuring head movement in dyads only), other studies may incorporate it as an experimental condition (e.g., measuring head movement in dyads versus triads). These decisions are reflected in the methodological approach used in conversation research, which will be further explored in the next section where I provide examples of measurements used to study conversation.

1.3.4 Conversation measurement approaches

As with any discipline, the methodologies employed in conversation research depend on the purpose of the study and the available resources. Approaches span a spectrum in terms of their level of constraint. Naturalistic methods involve observing interactions in real-world settings, capturing the spontaneity and authenticity of everyday conversations. Conversely, laboratory studies offer controlled environments to manipulate variables and isolate specific communication elements. Below (Table 1.3), I present a subset of examples illustrating the spectrum of approaches, ranging from highly naturalistic to more constrained methodologies.

Table 1.3 A summary of different approaches to studying conversation with a few examples.

	Method	**Authors**	**Description**
1	Eavesdropping	Moore (1992), Bischoping (1993)	These studies explored gender differences in conversations by writing down overheard conversations.
2	Passive Sensing	Eagle & Pentland (2006); Harari et al. (2017); Mulwa & Kucker (2024)	Using wearables or smartphone apps, this technology collects social interaction data without recording conversations. Meta-information about conversations is stored for examining variable fluctuations over time (e.g., length, group size estimation).
3	Naturalistic Studies	Cuperman & Ickes (2009); McFarland et al. (2013); Tschacher et al. (2014)	Recording interactions with minimal intervention in various contexts like speed dates, patient-therapist sessions, and negotiations. Reusing data for different purposes but caution is needed for contextual differences.
4	Ecological Momentary Assessments (EMA)	Burke & Naylor (2020); Intille et al. (2016); Schinkel-Bielefeld (2020)	Assessing certain real-life situations, including conversations, in participants' daily lives. The assessment can be based on the frequency of certain interactions or participants' perceptions.
5	Conversation Analysis	Goodwin & Heritage, (1990); Hepburn et al.,	A qualitative approach involving detailed transcriptions and annotations of conversation

		(2012); Mondada (2016); Sacks et al. (1974)	recordings to understand patterns in e.g., turn-taking, adjacency pairs, and repair.
6	Laboratory conversational studies	Hadley et al. (2021, 2019); Hadley & Culling (2022); Holler & Wilkin (2011); Petersen et al. (2022)	Using conversation prompts to elicit natural-like conversation to explore verbal and non-verbal communication behaviour such as speech, head movement, gestures, body posture.
7	Conversation Interventions	Epley & Schroeder (2014); Hirschi et al. (2023); Huang et al. (2017)	Studying how manipulations such as number of questions asked or response time, impact outcomes like likability and well-being in conversations.
8	Communication Games	Baker & Hazan (2011); Beechey et al. (2019); Fusaroli et al. (2012); Galantucci & Garrod (2011); Scott-Phillips et al. (2009)	Experiments using games to study components of communication. Allows precise predictions but may have limited generalizability beyond specific game contexts.
9	Just-follow-conversation tasks	Hygge et al. (1992); Shiell et al. (2023); Whitmer et al. (2024)	Recording volume preferences, eye movement or other behaviours while following a recorded conversation. This has been used to investigate turn-taking, understanding and sound perception.
10	Brain Recordings	Carlile & Keidser (2020); Drijvers & Holler (2022); Kang & Wheatley (2017); Zadbood et al. (2017)	Using brain imaging techniques like fMRI and dual EEG to explore brain dynamics during conversation, providing insights into comprehension, synchrony, and neural activity.

Each of the methodologies summarised in Table 1.3 has its advantages and limitations. Highly naturalistic methods, such as eavesdropping or passive sensing, provide information about real-world communication behaviours but

often at the expense of control over variables, which can lead to challenges in generalising findings across contexts. Laboratory studies and communication games allow for greater experimental precision, enabling researchers to isolate specific components of conversation. However, these methods may suffer from reduced ecological validity, as the highly controlled conditions may not fully replicate the spontaneity of everyday interactions. Qualitative approaches, such as conversation analysis, offer rich, detailed descriptions of conversational patterns but can be time-consuming and may be limited by subjective interpretation. Meanwhile, methods like ecological momentary assessments (EMA) strike a balance by capturing data in real-life settings, though they rely heavily on self-reporting, which can introduce bias. Then, the increased use of technology in approaches like passive sensing or brain recordings presents exciting opportunities for gaining objective data. Yet these methods may miss the interpersonal dynamics, reducing conversations to quantifiable metrics or neural signals that do not necessarily account for the subjective experiences of individuals. One possible solution to these limitations is combining multiple approaches to capture a more holistic view of conversation. For example, integrating subjective self-reports with observable, measurable behaviours could provide a deeper understanding of both the internal experiences of participants and the external manifestations of those experiences.

1.3.5 Processes that facilitate conversation

In the previous sections I have discussed that successful conversations involve information transfer, feedback, holistic integration with social experience, and have acknowledged that various different aspects of the conversation situation will affect perception of success. Next, I focus on how people adjust behaviour to increase conversation success.

Here, I turn to the cognitive, and neural mechanisms that could underlie successful conversation behaviours. In the following paragraphs, I explore some of these mechanisms. First, I discuss basic behaviour rules, followed by processes that involve responding to a communication partners' signals. Finally, I examine processes that involve monitoring the partner at a higher level and adjusting moment-to-moment behaviour. Each of these mechanisms

may contribute to conversation success. They could be used in different situations and could evolve as conversation duration increases or the interlocutors' needs are revealed.

1.3.5.1 Behaviour rules

One process that could lead to beneficial conversation behaviours is based on the fact that human behaviour follows certain rules (Hadley et al., 2022). These rules might deal with (but not limited to) face orientation, physical distance, and timing. Usually, people have in-person conversations facing each other. Although not strict, there is a universal understanding that the physical distance has different social functions. For instance, Hall (1966) defined the intimate space between 0-45 cm, the personal space 45-129 cm, the social space between 129-365 cm and the public space between 365-762 cm. Respecting an optimal interpersonal distance during conversations could facilitate access to verbal and non-verbal communication cues without triggering negative emotions (e.g., due to personal space invasion). Timing governs good coordination in conversations, and it appears to have an essential role in turn taking behaviour (Meyer, 2023; Stivers et al., 2009) as well as in mimicking non-verbal behaviours (Bekke et al., 2024).

1.3.5.2 Entrainment

Although referred to using different terms (alignment, adaptation, adjustment, mimicry, priming, coordination, and convergence), entrainment is known as the process through which conversation partners become more similar to each other over time (Chartrand & Bargh, 1999; Horton & Keysar, 1996; Levitan et al., 2012; Wynn et al., 2022). Within the context of interaction, there is evidence that this phenomenon appears at different levels. At the speech level, it has been shown that during conversation, communication partners become more similar in terms of acoustical characteristics such as fundamental frequency, voice quality and phonetics (Ostrand & Chodroff, 2021; Pardo, 2006), speaking rate (Levitan et al., 2012), but also linguistic style (Danescu-Niculescu-Mizil et al., 2011; Garrod & Pickering, 2004) and syntactic structure (Reitter et al., 2011). Similarity

between interlocutors has been observed as well in non-verbal behaviour such as gestures (Holler & Wilkin, 2011), facial expressions (Drimalla et al., 2019), and body position (Chartrand & Bargh, 1999; Paxton & Dale, 2017). Moreover, evidence suggests that, during a listening task, there is neural alignment in the cortical activity between the speaker and the listener (Stephens et al., 2010).

During conversation, participants are caught in a joint activity where they build flow and understanding using each other's facial expressions, words, and gestures. One explanation of this phenomenon is based on the existence of mirror neurons that trigger automatic imitation of others (Dijksterhuis & Bargh, 2001). This approach is known as the Interacting Alignment Theory and proposes that the entrainment of behaviour during interaction is automatic and unconscious (Garrod & Pickering, 2004). Another explanation for entrainment comes from the Grounding Theory (Clark & Wilkes-Gibbs, 1986) and highlights participants' conscious desire to establish a common ground. This is accomplished through dialogue where participants ensure a shared understanding of concepts. The theory rests on the assumption that every interaction is driven by a willingness to collaborate (Stalnaker, 2002). This is further supported by the Shared Reality Theory, which proposes the idea that people are consciously motivated to seek evidence that they have something in common with their conversation partners (Echterhoff et al., 2009). Furthermore, Metzing & Brennan, (2003) demonstrated that grounding is a partner-specific process. In situations where the conversation group expands, the concept must be re-established or abandoned to maintain effective communication. This has implications for the effectiveness of entrainment in group conversations, where finding similarity between participants might prove to be more difficult.

Automatic or not, being similar to a conversational partner has been shown to have various positive effects: increased liking among participants (Chartrand & Bargh, 1999), task success, coordinated turn-taking behaviour and fewer interruptions (Nenkova et al., 2008) and increased positive interactions in married couples (Lee et al., 2010).

1.3.5.3 Accommodation

Conversation partners are not always similar; at times, they may have diverse needs, backgrounds or communication styles. In such situations, establishing alignment or common ground might pose a challenge, and sustaining a successful conversation would necessitate a form of accommodation. For example, research has shown that people adapt their behaviour in various ways: they speak louder and use more gestures in noisy environments (e.g.: Lombard effect, (ISO, 2003)), adjust their speech speed and pause duration when addressing older audiences (Kemper, 1994), and modulate their speech level based on the degree of hearing loss of their conversational partner (Beechey et al., 2020). These adaptive behaviours find explanations in the Audience Design theory (Clark & Murphy, 1982), showing how behaviours are shaped to suit the context, and the Theory of Mind (Hamilton & Lind, 2016), which refers to our ability to consider the other person's perspective and adjust accordingly.

Although it is unclear how these mechanisms (behavior rules, entrainment, and accommodation) interrelate, it is likely that a conversation will be perceived as more successful when interlocutors respect each other's personal space, build conversation using each other's words, and accommodate each other's needs. One of the most common need for accommodation can arise from sensory impairments such as hearing loss. Since this thesis focuses on adults with hearing loss, the next section describes how hearing loss can impact conversation.

1.4 Background - hearing loss

"Blindness separates people from things; deafness separates people from people."

Helen Keller

1.4.1 Hearing loss

Hearing loss is increasingly affecting millions of individuals. In the UK it is estimated that approx. 1 in 6 adults is affected by some form of hearing loss (Vos et al., 2016). Hearing loss is a commonly used term in medical contexts to describe a measurable reduction in hearing ability ranging from mild impairment starting at 25 dB loss (for adults) to profound deafness. It can be congenital or acquired later in life. However, hearing difficulties can extend beyond what is detected by audiometric tests, as seen in conditions like hidden hearing loss and tinnitus. The term 'impaired hearing' is therefore often used to highlight the broader experience of hearing challenges. In this thesis, I refer to people with moderate hearing loss as 'people with impaired hearing'. Their specific hearing thresholds are defined in the methods sections of each study.

In the context of hearing disability classification (according to the International Classification of Functioning, Disability and Health (World Health Organization, 2001), four levels of hearing functioning are recognized. The lowest level is represented by 'hearing' which refers to the passive perception of sound (e.g., hearing a car passing by). It is followed by 'listening,' which refers to the process of hearing intently and attentively (e.g., listening to the weather forecast on the radio). 'Comprehending' refers to the process of understanding the information received. Lastly, the highest level of functionality is represented by 'communicating,' referring to the bi-directional transmission of information between two or more people (Kiessling et al., 2003). In this classification, the four levels are interconnected, and the proper functioning of each lower layer enhances the effectiveness of the superior layer (Figure 1.5). For example, listening (2nd layer) relies on the functioning of hearing (first layer). Similarly, effective communication (layer 4) depends on

the optimal functioning of the layers below. The effect of hearing loss can also be seen in the psychosocial functioning of people with hearing loss, which could be the fifth layer (Figure 1.5) (Dalton et al., 2003; Li et al., 2014; Strawbridge et al., 2000). Compared to individuals with normal hearing, individuals with hearing loss report lower rates of employability and levels of work-life satisfaction (for a review see Granberg & Gustafsson, 2021), retire earlier (Helvik et al., 2013), reduced perception of well-being and quality of life (Ciorba et al., 2012; Morgan et al., 2002; Scarinci et al., 2008), and even higher levels of anxiety and depression (Li et al., 2014).

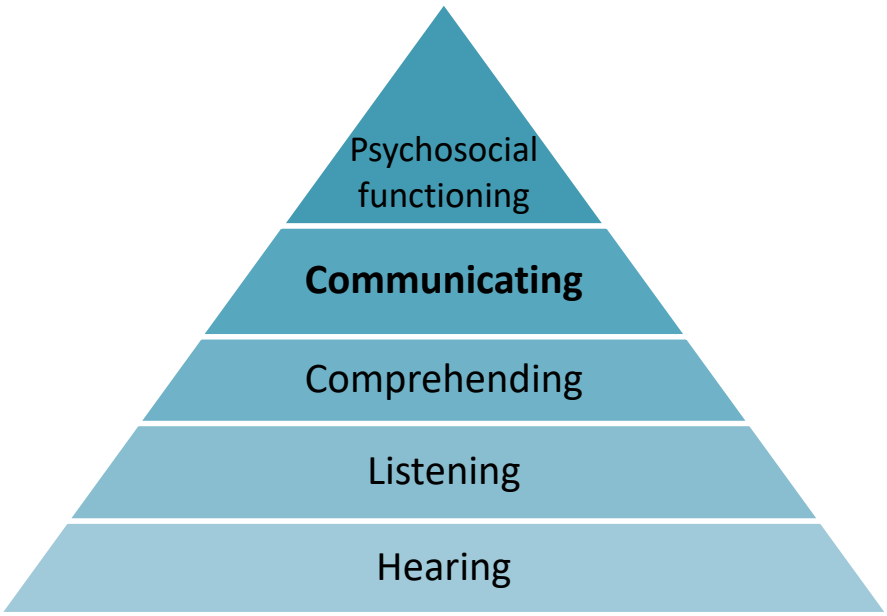


Figure 1.5 Pyramid chart showing the layers of hearing functioning adapted after ICF 2001 and Kiessling et al. 2003. The "communicating" layer shows how dependent conversation is on the inferior layers (hearing, listening and comprehending) and its importance for the superior layer (psychosocial functioning).

1.4.2 Hearing aids

Hearing aids are small wearable devices designed to amplify and modulate sound for individuals with hearing loss. These devices consist of a microphone to pick up sounds, a processor to increase the volume of the signals, a loudspeaker to deliver the amplified sounds to the ear, and a battery for power. Wearing hearing aids has been shown effective in improving psychosocial functioning in various ways. People with hearing loss seem to

experience reduced levels of anxiety and depression after being fitted with hearing aids (Joore et al., 2002), and improved quality of life (Abrams et al., 2005). The benefits derived from hearing devices extend beyond the users themselves, positively impacting those in their social circles (Scarinci et al., 2008). Research indicates that when individuals with hearing impairment use hearing aids, not only does their own quality of life improve, but there are also positive effects on the well-being of their frequent communication partners (Scarinci et al., 2008). Furthermore, the provision of hearing aids has been associated with a decrease in communication effort (Beechey et al., 2020) as well as a reduction in distress caused by communication difficulties for both individuals with hearing impairment and their partners (Brooks et al., 2001).

But while hearing aids can significantly improve the quality of life for some individuals with hearing loss, they do not completely restore normal hearing. The effectiveness of hearing aids can depend on factors such as the type and degree of hearing loss, individual preferences, and the proper fitting and adjustment of the device by a hearing care professional. Furthermore, communication difficulties do often remain.

1.4.3 Conversing with hearing loss

Regarding the consequences of hearing loss on communication behaviour, individuals with hearing loss face numerous challenges. Common difficulties include trouble understanding speech (Arlinger, 2003), particularly in noisy environments (Healy & Yoho, 2016), increased communication effort (Beechey et al., 2020), and the need for frequent repetitions (Lind, 2013; Lind et al., 2004, 2006). These challenges often result in an increased reliance on visual cues, such as lip-reading, and other cognitive compensatory processes to fill in gaps in speech (Rönnberg et al., 2019). Moreover, conversing in groups is one of the most reported difficulties of people with hearing loss (Vas et al., 2017). Following conversation, competing for holding the floor, and detecting who is talking without missing too much content might be more challenging for people with hearing loss. Social withdrawal, listening-related fatigue, and frustration can also emerge due to the struggles in comprehending spoken

communication (Palmer et al., 2019). Hence, individuals with hearing loss are more likely to avoid conversations in noisy settings compared to those without hearing impairments (Strawbridge et al., 2000).

Hearing aids can enhance the conversational process by improving speech intelligibility. However, their effectiveness appears to be limited in noisy environments, where individuals are surrounded by loud background noise (Healy & Yoho, 2016). The performance of hearing aids is often assessed in controlled listening situations (audiometric tests) or speech production paradigms, where individuals listen to words or statements and repeat what they hear. While these paradigms can provide valuable insights, unfortunately, they overlook the complexities that occur in conversations. There is a need to understand how hearing aids can be evaluated in realistic conversational scenarios (Smeds et al., 2020) and how they could contribute to conversation success.

Much of the research in communication within hearing sciences has focused on understanding difficulties, such as how individuals with and without hearing loss cope in challenging environments like high background noise. However, there has been little attention given to studying behaviours during successful conversations. It is essential to recognize that conversation success cannot be solely defined by the absence of difficulty. Successful conversation entails more than just the avoidance or management of challenges. It is possible to have difficult conversations that are successful and seemingly easy conversations that are considered unsuccessful. Therefore, in the next section I will focus on defining conversation success, considering factors beyond the mere absence of difficulty.

1.5 Conversation success

In previous sections, I explored aspects of communication and the conversation situation that could relate to conversation success. First, I looked at the structure of the communication process as described in different theoretical models of communication. I showed how these models fail to represent the complexity of real-life conversations. Then, I discussed how factors such as group size, goals, and familiarity can impact the perception of

conversation success. Finally, I provided examples of mechanisms that could contribute to beneficial behaviours and thus to success. But despite these insights, defining conversation success is unclear.

To the best of my knowledge, only one clear attempt has been made to define a concept close to conversation success and it warrants discussion. Spitzberg (2000) tackles the notion of "good conversation", offering a review (Table 1.4) of the criteria needed to assess the quality of a conversation. In his perspective, a good conversation is perceived when 1) the participant evaluates the conversation as having led to the achievement of his objectives, 2) the participants adhere to the rules of the situation and adapt to the context, 3) prior expectations were fulfilled for all participants resulting in positive affective response, 4) the outcomes achieved through the conversations were higher than the investment, 5) there was successful grounding between participants, and 6) a specific communicative task was achieved.

Table 1.4 A summary of the criteria involved in determining what constitutes a "good" or "bad" conversation according to Spitzberg (2000).

	Criteria	**Definition**
1	Effectiveness	The extent to which a communicator achieves his objectives.
2	Appropriateness	Conformity to the rules of a situation; perceived fitness or legitimacy of a communicator's behaviour in a context.
3	Satisfaction	Positive affective response to the fulfilment of previous expectancies.
4	Efficiency	Extent of valued outcomes achieved relative to the amount of investment.
5	Verisimilitude	The extent to which the symbol-referent link is well defined in the minds of the interpreters.
6	Task-Achievement	Refers to achieving specific communicative tasks or objectives.

While the criteria descrived seem plausible, the evaluation of such criteria becomes complex in practice. For instance, criteria 1 and 6 assume that all conversations have an objective. Some conversations lack a clear, conscious goal, and objectives may emerge or fade during the interaction. Then, appropriateness, while often considered important in conversations, may not

always be necessary for a conversation to be successful. In fact, in conversations involving interlcoutors from different cultural backgrounds, the absence of appropriateness can sometimes result in humorous reactions or responses. The criterion of satisfaction assumes that every conversation is conducted with prior expectations, which may not always be true. Criterion four, efficiency, assumes that conversation is a give-and-take exchange. Receiving more than giving implies success, while giving more than receiving implies lack of success. While this might be true in transactional conversations, it is not as clear-cut in relational conversations. The fifth criterion, verisimilitude, reflects the entrainment process described earlier. Based on the assumption that conversations typically involve a collaborative goal, the more alignment between interlocutors, the better the conversation is perceived to be.

Additionally, determining who decides if the criteria are met is not always obvious. Spitzberg (2000) considers that "good communication is located in perception rather than behaviour". He recognises three different sources of perception: ‘actor’ (the participant in the conversation), ‘coactor’ (the conversational partner of the ‘actor’) and ‘observer’ (an external bystander that did not take part in the communication process but followed it from outside). The studies involved in this thesis explore conversation success from all three perspectives.

1.6 Thesis aims

Despite the existence of numerous theories of human communication and well-documented processes that facilitate conversation, the precise factors that contribute to a successful conversation remain unknown. This lack of understanding is an obstacle to any efforts aimed at measuring conversational success. Clarifying the concept of successful conversation is essential, particularly for individuals with hearing loss, as it could pave the way for advancements in hearing aid technology and improve their overall communication experiences. In this thesis, I therefore attempt to conceptualise conversation success examining the perspectives of people with normal and impaired hearing. The conversation success investigation presented in Chapter 2 covers both one-to-one and group conversations. However, the following

chapters will focus exclusively on group conversations. Despite their prevalence in daily life, multi-party conversations have received little attention in the hearing science literature. Chapter 3 will shed light on the complexities of collecting data in four-way conversations, then Chapter 4 will investigate the self-reported perception of conversation success. Although existing studies on conversation employed different methodologies, none to my knowledge has linked communication behaviours to the self-perception of conversation success. Hence, Chapter 5 will investigate what behaviours are associated with conversation success. Finally, Chapter 6 will explore conversation success from third-party observers' perspective.

This thesis will investigate how individuals relate to the concept of conversation success in three contexts: (1) when they imagine an ideal conversation, (2) when they participate in successful and unsuccessful group conversations, and (3) when they watch other people having a group conversation. The seven chapters of the thesis revolve around three studies (figure 1.6).

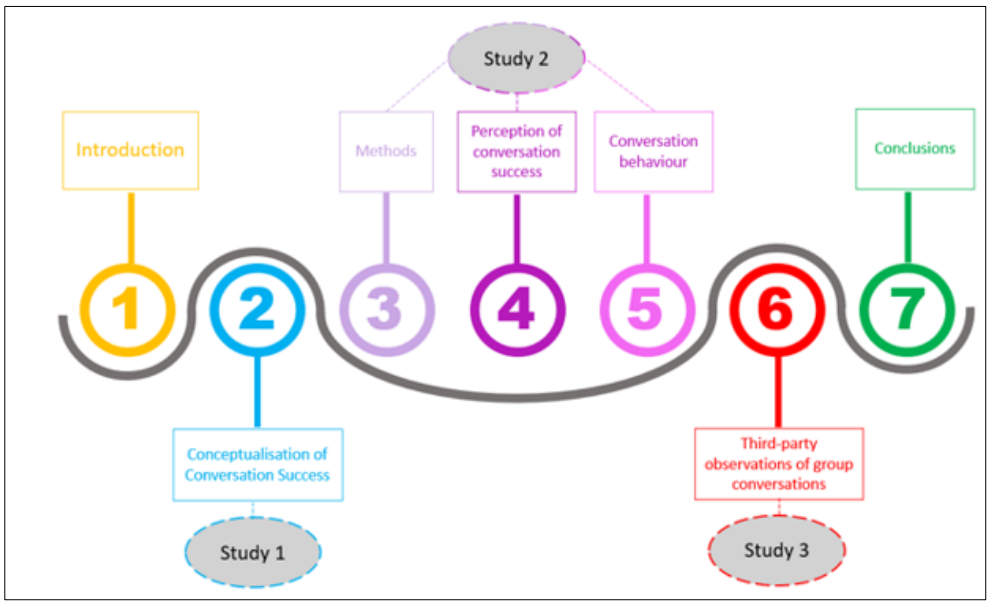


Figure 1.6 Diagram showing the sequence of the studies and their corresponding chapters

1.7 Glossary

Dyad – group of two people

Triad – group of three people

Quartet – group of four people

PwNH – person/people with normal hearing

PwIH – person/people with impaired hearing. In this thesis, all IH participants have moderate hearing loss.

HA – aided. It refers to Study 2 conversations in which participants with impaired hearing wore their hearing aids

UA – unaided. It refers to Study 2 conversations in which participants with impaired hearing did not wear their hearing aids.

GCM – group concept mapping

OPA – ordinal pattern analysis

PCC – statistical value used in OPA showing the percentage of correct clasifications observed in the collected data with respect to the hypothesis

AV – auditory-visual. It refers to Study 3 videos which have been presented with sound on

V – visual only. It refers to Study 3 videos which were presented without sound

Chapter 2

What is conversation success? (Study 1)

This study was published in the International Journal of Audiology. Raluca Nicoras, Sarah Gotowiec, Lauren V Hadley, Karolina Smeds & Graham Naylor (2023) Conversation success in one-to-one and group conversation: a group concept mapping study of adults with normal and impaired hearing, International Journal of Audiology, 62:9, 868-876, DOI: 10.1080/14992027.2022.2095538

The following chapter includes both previously unpublished and published figures. While several sections are rewritten, much of the wording in the methods, results and discussion are as published previously.

2.1 Chapter Summary

Although everybody wants to have successful conversations, the concept of conversation success is not clearly defined. This study aimed (1) to identify what constitutes conversation success and (2) to explore the importance of different components in one-to-one versus group conversations. The group concept mapping method was applied. Participants responded to two brainstorming prompts ("What does 'successful conversation' look like?" and "Think about a successful conversation you have taken part in. What aspects of that conversation contributed to its success?"). The resulting statements were sorted and rated in terms of importance for one-to-one and group conversation. Seven themes were identified: (1) *Being able to listen easily*; (2*) Being spoken to in a helpful way*; (3) *Being engaged and accepted*; (4) *Sharing information as desired*; (5) *Perceiving flowing and balanced interaction*; (6) *Feeling positive emotions*; (7) *Not having to engage coping mechanisms*. Three themes (1, 2, and 4) were more important in group than in one-to-one conversation. There were no differences in ratings of importance by hearing group.

2.2 Background

In the quest to quantify conversation success, it is imperative to start by establishing a clear conceptual framework. In other words, before measuring a phenomenon, it is necessary to define what the phenomenon is. This chapter lays the groundwork for the operationalization of conversation success, starting from the fundamental question: What constitutes 'conversation success'? This chapter will try to disentangle the nuances of this concept, paving the path toward an understanding of conversation success for individuals with diverse hearing abilities.

As previously seen in Chapter 1, communication theories share a common understanding of the conversational process, highlighting the exchange of information between a sender and a receiver through a communication channel, occurring within various contexts. These theories predominantly address the 'what' and 'how' of communication—what is being communicated (the message, the information, the meaning), and how it is being conveyed (the channel and required resources). However, these classic theories often overlook the 'why' aspect of communication. While the motivation behind communication may seem straightforward, we consider it important to explore this aspect, as it appears closely connected to the concept of success. If, in general, success implies the satisfaction of pre-existing goals, then conversation success might imply the satisfaction of pre-existing conversation goals that led to engagement in the communication process.

One of the most popular theories is that the fundamental purpose of human communication is to cooperate, hence everything involved in the process is or should be serving this purpose. Grice (1975) developed four maxims that are postulated to guide the participant in meeting the cooperation goal: 1) the maxim of quality – what is being said must be true; 2) the maxim of quantity – the contribution needs to be neither more nor less than what is required; 3) the maxim of relation – the contribution needs to be relevant; 4) the maxim of manner – the contribution needs to be clearly expressed, avoiding ambiguity and obscurity of expression. However, it is generally agreed that conversation also serves a more basic social purpose: being a means for

individuals to connect with others, foster relationships, convey affection, seek inclusion, and exercise control or persuasion (Schutz, 1966). This idea is further refined by Rubin et al. Rubin et al. (1988), who identified six motives underlying interpersonal communication, leading to the development of the Interpersonal Communication Motives scale. The six motives are (1) Pleasure, illustrating how individuals engage in conversations for enjoyment, stimulation, and inspiration; (2) Affection, showing how people use conversations as a platform to express care and appreciation for others; (3) Inclusion, highlighting how conversations can serve as a remedy for loneliness, enabling individuals to overcome social isolation; (4) Escape, revealing that people may engage in conversations to evade other activities or simply to avoid boredom; (5) Relaxation, illustrating how individuals harness conversations as a method to unwind and alleviate stress and (6) Expressing control, illustrating how conversations can be used as a persuasive instrument, allowing individuals to exert influence. More recently, Rubin's findings have been confirmed by Jensen (2018) who found that the primary reasons for communication include sharing and increasing enjoyment as well as creating and strengthening social bonds and that secondarily communication serves purposes associated with information exchange, including staying informed, facilitating collaboration, and sharing personal experiences. As seen in Chapter 1, such conversational goals are not mutually exclusive. People can pursue multiple goals at the same time and even change their intentions during the communication process (Yeomans et al., 2022).

While the literature does not provide a clear representation of what conversation success is, researchers have long discussed related concepts such as communication satisfaction (Hecht, 1978) and attributes of good conversations (See et al., 2019). Hecht (1978) defines 'communication satisfaction' as the positive emotion experienced following successful and fulfilling communicative interactions. However, this definition relies on synonyms or constructs that are themselves poorly defined. If satisfaction is the emotion that follows a successful and fulfilling conversation, it prompts the

essential question of what precisely constitutes a successful and fulfilling conversation.

Clark et al. (2019) investigated the key attributes that people value in conversations, identifying four central themes: mutual understanding and shared perspectives, trustworthiness, active listening, and humour. The theme of ‘mutual understanding and common ground’ highlights the importance of understanding not only the words spoken but also the underlying intent and meaning. ‘Trustworthiness’ is shown to be a necessary attribute for a good conversation, and it is seen as an opening for more personal conversations. ‘Active listening’, another attribute of a good conversation, refers to the ability of the interlocutors to show engagement and willingness to participate in conversation. Lastly, 'humour' surfaces as an essential attribute, highlighting the role of amusement in shaping perceptions of conversation.

Crafting a universally applicable definition of conversation success might prove challenging, given the variability dictated by context, motivation and participants involved. For instance, while in business and management conversations success may revolve around productive feedback sessions and efficient decision-making, in the field of psychotherapy it could be associated with the meaningful verbalization of emotional experiences. Everybody wants to have successful conversations; however, our focus remains on older adults with hearing loss having casual everyday conversations in small groups (Chapter 1). In hearing sciences, hearing aid manufacturers are committed to designing devices that compensate for hearing loss, aiming to restore people’s desire to fully engage in their social world. Knowing what conversation success is could facilitate faster progress, by ensuring that efforts are focused on enhancing functionalities that have the most importance for users of hearing devices.

This study aims to address this knowledge gap by employing Group Concept Mapping (GCM), a mixed-method participatory research process (Trochim, 1989), to collect insights from adults with both normal and impaired hearing. The research focuses on two key questions: (1) What constitutes ‘conversation success’? and (2) How do the components of ‘conversation

success’ differ in importance between one-on-one and group conversations? GCM blends qualitative and quantitative research in a sequential process that actively involves participants. In addition to sharing their perspectives on the topic, participants are engaged in grouping ideas and rating them on various scales (e.g., importance in one-on-one conversations and importance in group conversations), thereby minimizing the potential influence of researchers' bias and assumptions.

Traditionally, GCM has been employed in business management (Trochim et al., 1994), but in recent years, it has gained popularity in healthcare research, often exploring patients' needs (Zevon et al., 2007). In audiology research, this method has been used to examine factors in client-clinician interaction that may influence hearing aid adoption (Poost-Foroosh et al., 2011), to investigate strategies for mitigating the impact of chronic tinnitus on children and adolescents (Bennett et al., 2018), to explore the use of tele-audiology in remote hearing-aid support services (Glista et al., 2020), and to gather perspectives from hearing-aid users, their significant others, and hearing care professionals on how eHealth can meet patients' needs (Meyer et al., 2021). Our GCM process involved two brainstorming prompts to probe the concept of conversation success. This strategy, as previously demonstrated in conceptualizing 'good health' (McCaffrey et al., 2018), proves effective in addressing the broad nature of 'conversation success.' The next section describes how GCM was implemented in this study.

2.3 Materials and methods

Participants were recruited from the Hearing Sciences Scottish Section Participant Database, and each gave their written consent for participation in the study. This research has received ethical approval from the West of Scotland Research Ethics Committee (18/WS/0007) and the NHS Greater Glasgow and Clyde R&D (GN18EN094). Given the COVID-19 pandemic, the entire study unfolded online. Participants were compensated for their time with an Amazon voucher of £10/session.

2.3.1 Participants

People with normal (PwNH) and impaired (PwIH) hearing were invited to participate. They were selected based on an audiogram and demographic data stored in the Hearing Sciences Scottish Section (HSSS) participant database. Invited participants were 50-75 years old and were fluent in English.
Inclusion criteria for PwNH:

- Age 50-75 years
- Better-ear pure-tone average hearing threshold less than 20 dB HL
- No threshold above 40 dB HL at any frequency
- Asymmetry of average thresholds maximum 10 dB HL.
- Fluency in English

Inclusion criteria for PwIH

- Age 50-75 years
- Better-ear pure-tone average hearing threshold between 41 and 60 dB HL
- Asymmetry of average thresholds maximum 15 dB HL.
- Fluency in English

Hearing aids are often recommended for people with hearing losses in this range, thus although it was not an inclusion criterion, all participants with hearing loss were also hearing aid users.

Thirty-five participants (N = 16 PwNH, N = 19 PwIH), approximately age-matched (mean PwNH = 62, mean PwIH = 65) agreed to participate in the first activity (brainstorming), 24 (PwNH =10, PwIH = 14) in the second activity (sorting the statements into groups), and 29 (PwNH = 13, PwIH = 16) in the third activity (rating the statements). 27 participants completed all activities, 29 completed two of the three activities and 6 completed only one activity The GCM method does not impose a specific number of participants per activity, but Trochim (1989) recommends groups of at least 10 participants to ensure diversity of ideas. Table 2.1 summarises the demographic data and participation rates in the different GCM activities.

Table 2.1 Demographics and participation rates

Age in years, mean (SD)	63.9 (6.5)	64.5 (6.4)	64.3 (6.4)
Activity	**Brainstorming**	**Sorting**	**Rating**
Gender			
Female	19	12	16
Male	16	12	13
Hearing condition			
Normal hearing	16	10	13
Impaired hearing	19	14	16
Work status			
Retired	20	14	17
Paid employment	10	7	7
Not in paid employment	5	3	5
Total	**35**	**24**	**29**

2.3.2 Group concept mapping steps

The GCM method was implemented as outlined below. Statements coming from all participants, both PwNH and PwIH, were processed, sorted, and rated together in order to obtain only one concept map that covered the views gathered from both groups.

(1) Brainstorming: Online individual unsupervised brainstorming sessions were conducted via an online survey platform (JISC, www.onlinesurveys.ac.uk). Participants were prompted to generate statements by answering the following questions: "What does 'successful conversation' look like?" and "Think about a successful conversation you have taken part in. What aspects of that conversation contributed to its success?". The first question is more general and elicits participants' imagination in thinking about conversation success. The second question is more specific and explores the post-hoc perception of conversation success by eliciting participants' memory. Participants were instructed (Appendix A) to write a minimum of one statement and a maximum of ten statements for each question. In addition to the brainstorming prompts, participants answered

demographic questions related to their work status, and whether they actively use hearing aids.

Next, the researchers condensed the brainstorming statements to create a logical set of statements (Trochim, 1989). The statements were trimmed and refined using the following strategies: removing duplicates, splitting items that express two or more topics in one statement, editing jargon and personal information, and excluding statements that did not answer the brainstorming prompt. The final list of statements was reworded for clarity.

(2) Sorting the statements: This activity was carried out by participants using the card-sorting functionality provided on www.provenbyusers.com, an open-source research platform. Participants who took part in this activity were given a set of virtual cards. One refined statement from the brainstorming step was written on each card. Participants were asked to organise the cards into groups of similar ideas in a way that made sense to them (instructions in Appendix A). They were told there was no right or wrong answer and asked to create at least five groups and to label each group with a unique title. A card could be put alone in a group if it was judged to stand alone as a unique idea, unrelated to the other cards. Participants were also asked not to create a ’miscellaneous’ or ‘other’ group. Three participants did not follow instructions and created less than five groups (n=2) or created random group names A, B, C (n=1). These participants were excluded from the analysis; therefore, they are excluded from Table 1.

(3) Rating the statements: In the rating activity, participants were asked to rate how important each aspect of success (i.e.: each statement) was to them. For this activity, we again used the JISC platform (www.onlinesurveys.ac.uk). Participants rated the statements on a five-point Likert scale (1= not important at all, 2 = somewhat important, 3 =

moderately important, 4 = very important, 5 = extremely important) in A) 1-to-1 conversation and in B) group conversation. Four statements referred to hearing-aid use (e.g.: Not needing to constantly adjust the volume on my hearing aid.) therefore a N/A (not applicable) response option was added. This step concluded the participants' involvement.

(4) Data analysis and interpretation: The data were analysed using R-CMap Software, an open-source tool implemented in R (Bar and Mentch 2017). The software follows the steps of the GCM as developed by Trochim (1989). Sorting data is aggregated, and a similarity matrix is obtained. Data is analysed using multidimensional scaling (MDS) analysis, producing x and y coordinates for each statement, corresponding to a point on a map (Figure 2.1.). The extent to which the original relative distances are preserved in the two-dimensional space is measured by a "stress" statistic. The more the MDS algorithm successfully preserved pairwise distances, the lower the stress (Bar & Mentch, 2017).

R-CMap software then applies a hierarchical cluster analysis to the MDS output, grouping these points into clusters. Each point represents one statement, and each cluster of points represents a theme. The result is a set of themes, where themes are maximally distinct (i.e., distant) from each other, and the points within each theme are maximally similar (i.e., close) to one another. Finally, the analysis computes averages of importance for each type of conversation across participants to arrive at a statement average and then computes averages across all statements within a theme to arrive at a theme average on the scale in question. R-Cmap's analysis of variance (ANOVA) and post-hoc analysis (Tukey) are limited to means comparisons between themes within participant groups and within the type of conversation. Therefore, additional two-way mixed ANOVA analysis to compare means between NH versus IH groups in 1-to-1 and group conversations was computed using SPSS. The reliability of sorting and rating data

was obtained by applying a split-half correlation process (Rosas & Kane, 2012; Trochim et al., 1994) in SPSS.

Interpreting the maps: Once the concept map was obtained, the researchers analysed, interpreted, and named the themes. The names were given considering the content of each theme and were inspired by the titles participants gave to their groupings in the sorting step.

2.4 Results

The outcomes for each activity are outlined below. We first report the results of the brainstorming, and then describe the concept map of themes generated from the sorting (clustering) activity. We then report the data resulting from the rating activity, which is split into multiple sections. We compare the importance of each theme by conversation type and by participant group, then we explore in-depth the differences by hearing group, and finally, we report differences between themes within the same type of conversation.

2.4.1 Brainstorming

Participants created 263 statements that were reduced to 73 after trimming (Appendix B). Each statement expressed an idea of conversation success and is represented by a number in Figure 2.1.

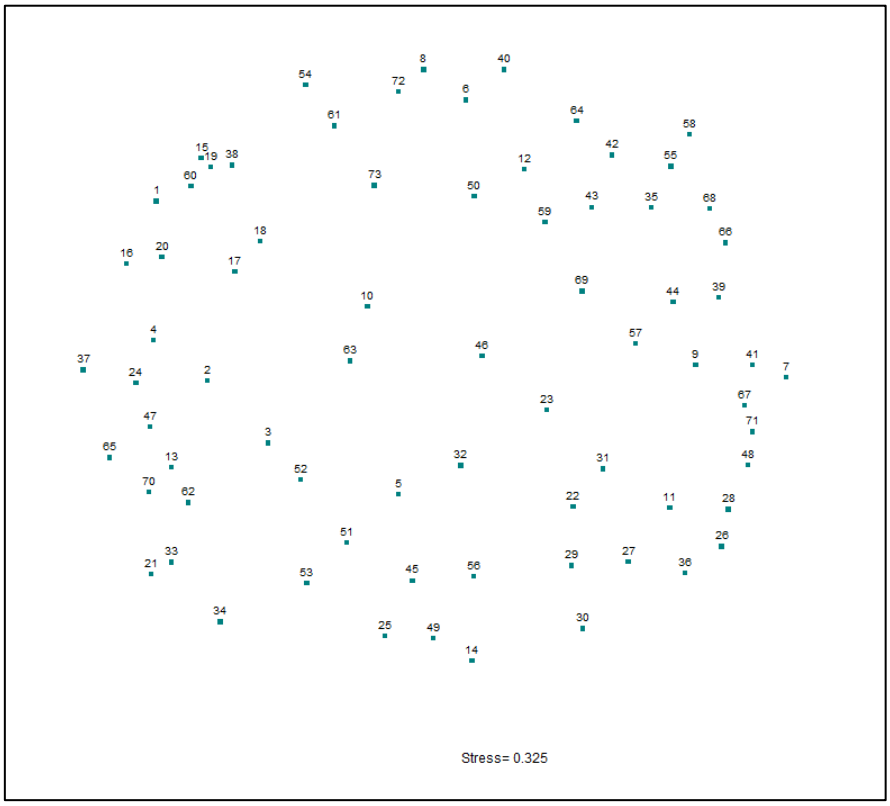


Figure 2.1 Point map showing the distances between the statements prior to applying the hierarchical cluster analysis, which is presented in Figure 2.2.

2.4.2 Sorting (clustering)

Participants created on average 6.3 groups (Min=5; Max=11; SD=1.62). MDS analysis resulted in the point map shown in Figure 2.1. Each point on the map represents one statement. The smaller the distance between two points, the more often the corresponding statements were grouped together by participants. E.g.: statement 1 (‘The speaker has a loud and clear speaking voice.’) was more often grouped together with statement 15 (‘The speaker makes their points succinctly, without rambling.’) than it was grouped with statement 30 (‘Feeling useful and appreciated’). This map has a stress index of 0.325 which is considered to be a value within the acceptable range, as a stress value below 0.369 (Trochim, 1989) indicates that discrepancies between the distances on the map and the values in the input similarity sort matrix are small.

2.4.3 Themes of conversation success

To identify the most interpretable division of the data into themes, we started from the minimum number created by participants (five), then increased the number of themes until all were internally consistent while being distinct from each other. As a concept map of six themes contained one theme with very dissimilar statements, and a map of eight themes contained two very similar themes, this left us with seven themes. As a supplementary check, we also analysed the content of the themes in a concept map version of three, four, nine, ten and eleven themes. In all options below seven the content of each theme lost homogeneity, and in all options above seven, the content of each theme lost uniqueness. The reliability of the concept map was analysed by randomly distributing the participants into halves and obtaining a concept map for each (Rosas & Kane, 2012; Trochim et al., 1994). Stress values obtained from split-half similarity matrices were correlated resulting in a Cronbach's alpha of .89. This is higher than .86, the average found in sixty-nine GCM studies (Rosas & Kane, 2012), indicating that the seven-theme map (Figure 2.2) possesses strong consistency.

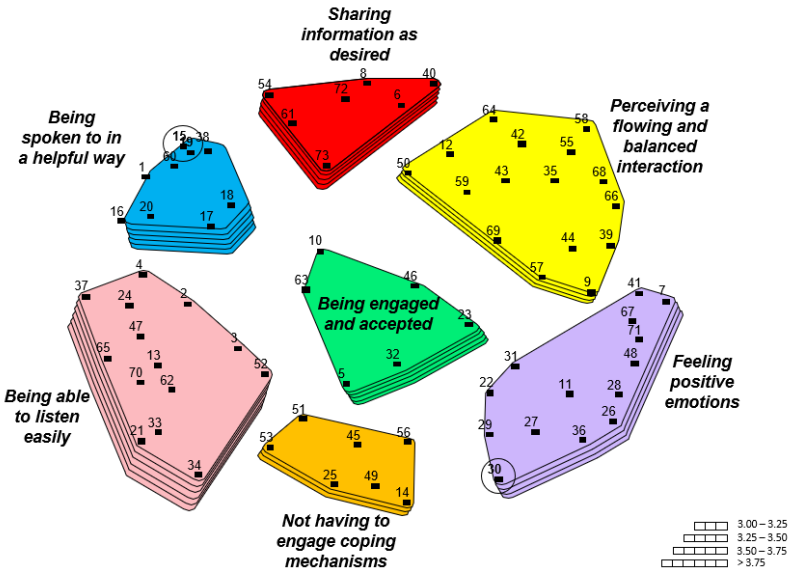


Figure 2.2 The map depicting seven themes of conversation success. The number of layers represents the ratings of overall importance given to each theme.

The seven themes related to conversation success are: (1) *Being able to listen easily*; (2) *Being spoken to in a helpful way*; (3) *Being engaged and accepted*; (4) *Sharing information as desired*; (5) *Feeling positive emotions*; (6) *Perceiving flowing and balanced interaction*; (7) *Not having to engage coping mechanisms*. The title of each cluster was inspired by the titles given by the participants and represents an umbrella for the statements inside. A short description of each theme along with some examples of statements can be seen in table 2.2. The full list of statements/theme is provided in Appendix B.

Table 2.2 Seven themes representing constructs of conversation success, a brief description, and examples of statements for each theme.

Themes (number of statements)	Short description	Examples of statements
Being able to listen easily (14)	Elements that can affect the listening process: background noise, working hearing aids.	Not having to strain to hear the other person(s). Hearing aids working (e.g.: not whistling, coming loose or batteries going flat). No distractions in the background.
Being spoken to in a helpful way (9)	Different attributes and responsibilities expected from the speaker.	The speaker has a loud and clear speaking voice. The speaker is not making too many points in one statement. The speaker talks fluently.
Being engaged and accepted (6)	Creating a connection between participants.	The listener shows interest in what I'm talking about. Feeling that you have learned something new at the end of the conversation. Body language reflecting engagement, e.g.: nodding, smiling.
Sharing information as desired (7)	Information exchange in task-oriented conversations, achieving outcomes.	Passing and receiving information. Communicating a want or task. Achieving a desired outcome.
Perceiving flowing and	Mutual engagement and maintaining	Participants don't interrupt or talk over the top of each other.

balanced interaction (16)	conversational dynamic.	A balance between asking questions and answering them. All participants contribute equally.
Feeling positive emotions (14)	Feeling good during and after the conversation.	Feeling joy in chatting with another human being. Participants are laughing and being funny. Leaving a conversation feeling inspired.
Not having to engage coping mechanisms (7)	Avoidance of negative feelings; finding comfort in the surroundings.	Not feeling anxious. Putting in minimal effort on my part. Not finding myself withdrawing from the conversation.

2.4.4 Ratings of importance

2.4.4.1 Statement level

Taken individually, the highest rated statement for importance in 1-to-1 conversations is statement 24# 'Hearing clearly, without having to assume what other the person was saying', part of the theme *Being able to listen easily.* The lowest rated statement is 41# 'Participants are laughing and being funny', part of the theme *Feeling positive emotions*. In group conversations, the highest rated statement is 33# 'Hearing aids working (e.g.: not whistling, coming loose or batteries going flat)', part of the theme *Being able to listen easily*. The lowest rated is the statement 57# 'Having some kind of familiarity or friendship between the participants.', part of the theme *Perceiving flowing and balanced interaction*.

2.4.4.2 Theme level

The average importance ratings of the statements within each theme ranged from 4.4 for *Being able to listen easily* (the most important) to 3.3 for *Not having to engage coping mechanisms* (the least important).

2.4.5 Themes importance in group versus 1-to-1 conversation

While the data obtained from the brainstorming and sorting activities were not specific to conversation type, the rating activity captured the importance

perceived by participants in 1-to-1 conversations and group conversations separately. Split-half reliability tests found a Spearman-Brown correlation above 0.90 for both types of conversation, consistent with the average correlation in prior GCM work (Rosas & Kane, 2012).

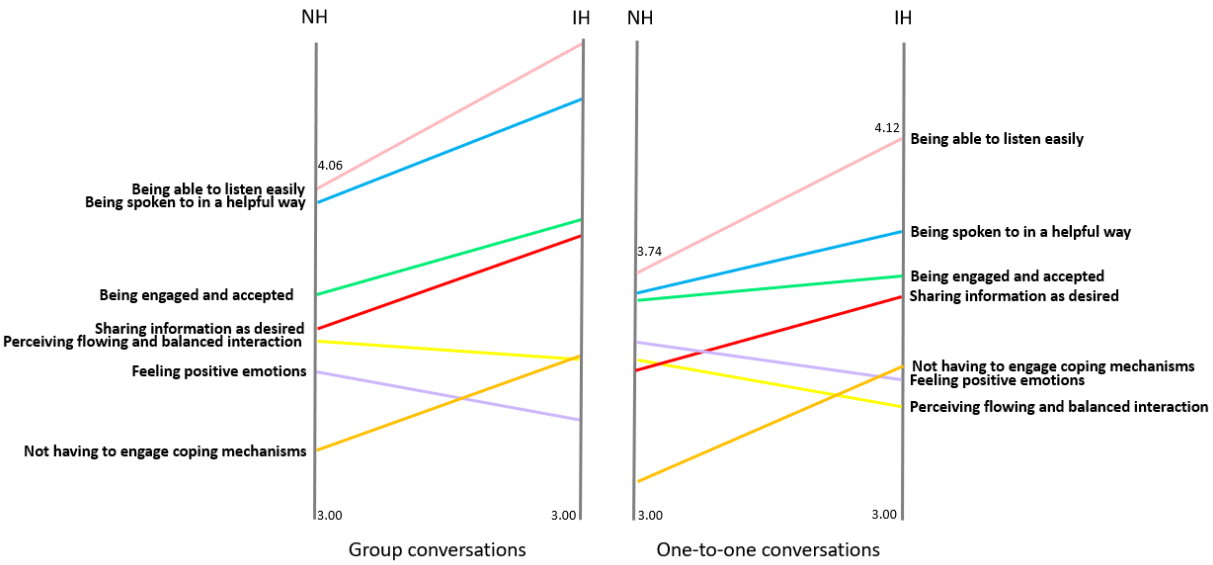


Figure 2.3 Laddergram showing differences between themes' average importance ratings per type of conversation for participants with normal and impaired hearing.

Seven two-way mixed ANOVA were performed, one for each theme, to analyse the effect of hearing ability (PwNH and PwIH) and type of conversation (1-to-1 and group conversation) on the importance given to each theme. Simple main effects results showed that the following three themes were statistically more important in group conversation than in 1-to-1 conversation: *Being able to listen easily* $F(1, 27) = 23.67$, $p = < 0.001$, *Being spoken to in a helpful way* $F(1, 27) = 13.52$, $p = < 0.05$ and *Sharing information as desired* $F(1, 27) = 4.40$, $p = < 0.05$. (Figure 2.3).

2.4.5.1 Differences between themes within the same type of conversation

ANOVA repeated measures analysis was used to compare the importance of each theme with all other themes for the same type of conversation, aggregating across participant groups. A small p-value indicates that at least one theme has a mean rating which is significantly different. Results showed that mean importance ratings differed between themes within 1-to-1 conversation $F(6, 66) = 3.24$, $p = 0.007$ and within group conversation $F(6, 66) = 9.07$, $p = <0.001$. Given these results, Tukey's method was used to perform pairwise comparison between all possible pairs of themes. Bonferroni correction for multiple comparisons was applied. *Being able to listen easily* was significantly more important than several other themes in both 1-to-1 and group conversation, whereas *Being spoken to in a helpful way* was only significantly more important than other themes in group conversations (see Table 2.3).

2.4.5.2 Exploratory analysis by hearing group

While the already mentioned two-way mixed ANOVAs did not indicate significant differences between hearing groups, in further exploratory analyses a marginal effect was found between PwNH and PwIH for the theme *Being able to listen easily* ($p = .059$) in 1-to-1 conversations, with higher scores of importance given by people with hearing impairment. Furthermore, while for PwNH only two themes (*Being spoken to in a helpful way* and *Being able to listen easily*) were rated to be significantly more important in group conversation than in 1-to-1 conversation, for PwIH four themes registered a significant increase in group conversation (*Being spoken to in a helpful way, Being able to listen easily, Being engaged and accepted*, *Sharing information as desired*). This is consistent with the notion that relative to people with normal hearing, people with hearing loss experience a greater contrast in difficulty between 1-to-1 and group conversations.

Table 2.3 Tukey p values for significant differences between themes for each conversation type

Theme contrast	Importance in 1-to-1 conversations p value	Importance in group conversations p value
Being able to listen easily - Not having to engage coping mechanisms	0.026	< 0.001
Being able to listen easily - Perceiving flowing and balanced interaction'	0.015	< 0.001
Being able to listen easily - Feeling positive emotions		< 0.001
Being able to listen easily- Sharing information as desired		0.036
Being spoken to in a helpful way - Not having to engage coping mechanisms		0.010
Being spoken to in a helpful way - Feeling positive emotions		0.001
Being spoken to in a helpful way - Perceiving flowing and balanced interaction		0.008

2.5 Discussion

The aim of this study was to identify the constructs related to conversation success as expressed by individuals with both normal and impaired hearing, and to explore their importance in 1-to-1 conversations vs group conversations. Our investigation revealed seven themes indicating distinct constructs of conversation success. The themes, ranked in order of their overall importance across conversation type and participant group are as follows: (1) *Being able to listen easily*; (2) *Being spoken to in a helpful way*; (3) *Being engaged and accepted*; (4) *Sharing information as desired*; (5) *Perceiving flowing and balanced interaction*; (6) *Feeling positive emotions*; (7) *Not having to engage coping mechanisms*. The themes *Being able to listen easily, Being spoken to in a helpful way* and *Sharing information as desired* were considered significantly more important in group conversations compared to 1-to-1 conversations across participant groups (PwNH and PwIH).

2.5.1.1 Themes and their content

Interestingly, some of our themes align closely with prior work (Rubin et al., 1988), while others are novel. The motives to engage in communication found by Rubin et al. (1988) are well reflected in the content of some themes. For Rubin et al., motivations to engage in interactions were inclusion, pleasure, affection, escape, relaxation, and control. Rubin's 'inclusion' involves sharing

feelings with others and overcoming loneliness, while ‘affection’ highlights the will to express caring and appreciation for others. These two motives are compatible with our third theme *Being engaged and accepted*. Furthermore, Rubin’s ‘pleasure’ motive shows that people communicate for amusement, entertainment, and ‘relaxation’ motive refers to using interaction as a method to unwind. These two motives correspond well to the content of our theme *Feeling positive emotions*. Additionally, this theme includes statements like 71# ‘Sharing a sense of humour and/or irony’ which confirms Clarke et al. (2019)’s results showing that humour is one of the most important attributes for a good conversation. Finally, Rubin's 'Control' was characterized by the use of communication as a tool to secure compliance, a notion that closely resonates with our theme *Sharing information as desired*. In the end, Rubin's 'Escape' motive did not find a direct parallel within our themes.

The theme rated most important for conversation success by both participant groups was *Being able to listen easily*. This theme contained items that reflect minimal effort e.g.: 13# ‘Not having to strain to hear the other person, with items indicating a quiet acoustic environment’, 65# ‘Being in a quiet place without background noise or music’, as well as with items expressing the necessity for well-functioning hearing aids e.g.: 33# ‘Hearing aids working (e.g.: not whistling, coming loose or batteries going flat)’. What these items share is their association with sound perception. Listening effort increases in various situations, such as a noisy environment, or when the listener has hearing loss, and this theme seems to integrate items that contribute to the effort involved in listening. For individuals with hearing loss, working hearing aids can alleviate effort (Ohlenforst et al. 2017; Picou, Ricketts, and Hornsby 2013), explaining their appearance in this theme.

The second most salient theme according to the participants’ ratings, *Being spoken to in a helpful way*, contains items that encompass a set of qualities expected from the speaker. For instance, elements such as voice volume and intelligibility are exemplified by statement 1# ‘The speaker has a loud and clear speaking voice’. In line with the theory of audience design (Clark and Murphy 1982), speakers who tailor their messages to the listener’s

needs are perceived as ‘helpful’ by listeners. This adjustment is consistent with the Lombard effect, whereby speakers increase voice intensity in noisy environments (Brumm & Zollinger, 2011). However, domains other than the acoustical can also contribute to speaking in a helpful way such as the way in which ideas are structured (19# ‘The speaker is not making too many points in the one statement.’) as well as other multimodal aspects of speech (18# ‘The speaker is looking towards the listener rather than around.’). These statements show that consistent information is perceived as helpful even when it is coming from different sources (auditory-visual), consistent with studies showing that multimodal speaker messages tend to be processed more swiftly than unimodal (auditory-only) messages (Holler & Levinson, 2019).

It is worth noting that both of the first two themes (*Being able to listen easily* and *Being spoken to in a helpful way*) are both concerned with ease of speech recognition. The former, *Being able to listen easily*, combines elements related to both the 'channel' (environment and devices) and the 'receiver' (hearing ability). In contrast the latter theme, *Being spoken to in a helpful way*, predominantly encompasses elements associated with the 'transmitter' or speaker component within the communication chain.

The *Being engaged and accepted* theme reflects the human desire for connection through interaction. In contrast to the previously mentioned themes, this theme reflects the psychosocial dimensions of conversation success. Individuals perceive conversation success when they experience a connection with their interlocutors on different levels: behavioural (5# ‘Body language reflecting engagement, e.g.: nodding, smiling.’), cognitive (10# ‘Being able to fully understand the topic.’), and emotional (32# ‘Feeling comfortable when asking or being asked for something to be repeated.’). From an evolutionary standpoint, the inherent human inclination to seek connection and belonging within communities aligns with this theme (Hall & Davis, 2017). The universal desire for successful conversations serves as a means to foster, expand, or strengthen social bonds.

Within the fourth theme *Sharing information as desired* statements refer to the goal of the conversation: 8# ‘Communicating a want or a task’

involving items that indicated the context of the conversation: 54# ‘Being able to clearly communicate with professionals (e.g.: doctor)’ as well as with items that were representing feelings associated with a transactional conversation: 61# ‘Feeling professional’. Thus whereas theme 3 is focused on relational conversations, the statements of theme 4 are particularly relevant for high informational conversations where success is perceived when the goal is achieved (Clark et al. 2019, Yeomans et al., 2022).

The theme *Perceiving flowing and balanced interaction* confirms that successful conversation is a dynamic interaction and requires active engagement from all parties (e.g.: 42# ‘Participants are fully engaged and contributing.’) as well as adherence to the unwritten rules of turn-taking (e.g.: 43# ‘Participants don’t interrupt or talk over the top of each other’) (Grice, 1975). In successful conversations, interlocutors resemble dancers that follow a certain rhythm, maintaining synchrony and balance while smoothly coordinating successive elements (Littlejohn and Foss, 2011).

Feeling positive emotions contains fourteen statements expressing feelings that can be experienced during a conversation (28# ‘Feeling joy in chatting with another human being.’) and continue or appear after a conversation (27# ‘Leaving a conversation feeling pepped up’). It is known that people want to converse because it is fun and stimulating (Rubin et al. 1988, Clarke et al. 2019), but what is noteworthy, is that our participants reflected on the continuation of positive feelings after the event. Sometimes the positive feelings related to the conversation can continue long after the conversation finished (e.g.: when partners are sharing their feelings of love for each other for the first time or when interlocutors are planning a holiday).

Finally, the theme *Not having to engage coping mechanisms* sheds light on the emotional challenges that can appear when trying to engage in or maintain successful conversations. Conversations, particularly for individuals with hearing impairments, are not always easy. For instance, during conversations they can feel negative emotions (e.g.: frustration, withdrawal) linked to hearing loss (Holman et al., 2023). Statements like 14# ‘Not feeling anxious’ and 49# ‘Not finding myself withdrawing from the conversation.’

show that people not only identify the positive feelings during conversation, but also the absence of negative feelings/coping behaviours.

These findings also show that conversation success is a shared responsibility among interlocutors. While some results (*Being able to listen easily*) show support for hearing aid use, others (*Being spoken to in a helpful way* and *Not having to engage coping mechanisms*) show support for a more holistic approach. In addition to hearing aid use, these results uphold the provision of family centred-care, participation in aural rehabilitation programs with frequent communication partners and even cognitive behavioural therapy.

Important to acknowledge is that although each theme is distinct, there are overlapping ideas between them. For example, under the theme *Perceiving flowing and balanced interaction* the statement 9# *'*A balance between seriousness and humour' is very close to the theme *Feeling positive emotions*. This conceptual overlap is reflected in the physical distance that appears on the map (Figure 2) between the statement 9 and the theme *Feeling positive emotions*. These overlaps are common in GCM (group concept mapping) studies since participants are encouraged to sort the statements in a way that makes sense to them, without being guided to think about a certain logic. Also, even if participants all answer the same brainstorming prompt, contradictory ideas can appear, e.g.: 67# 'Not having any pressure regarding the outcomes of the conversation.' and 72# 'Achieving a desired outcome'.

2.5.2 Ratings of importance per theme

Even in the case of the theme rated as the least significant, *Not having to engage coping mechanisms*, its average importance rating still exceeded 3 out of 5, marking it as 'moderately important.' As a result, none of the themes can be dismissed as lacking in importance. This is a natural consequence of the elicitation approach in the brainstorming step. While participants were not explicitly directed to articulate aspects of conversation success that they considered significant, it is plausible that the ideas spontaneously generated during this phase naturally reflected what they personally considered important.

Both *Being able to listen easily* and *Being spoken to in a helpful way* received the highest importance ratings across both types of conversation. These two themes are closely associated with the sensory facets of in-person communication, particularly hearing and vision. High importance ratings might suggest that sensory deficits come with a high cost for conversation success. For instance, a face-to-face conversation might go on and be perceived as successful even if it is not flowing smoothly, whereas a face-to-face conversation in which interlocutors lack the ability to hear or see each other makes it almost impossible.

Among the seven themes, six displayed a higher average importance rating in group conversations compared to 1-to-1 interactions. This observation implies that the challenges associated with conversation success tend to intensify as the number of participants increases. One plausible explanation for this phenomenon lies in the potential connection between importance and complexity. *Being able to listen easily*, *Being spoken to in a helpful way* and *Sharing information as desired* might pose greater difficulties when sharing time and space with multiple individuals, which could explain their higher importance ratings. Notably, this contrast in importance between 1-to-1 and group conversations was more pronounced for individuals with hearing impairments, suggesting that they exert greater effort when confronted with the simultaneous presence of multiple conversational partners. Conversely, the theme *Feeling positive emotions* was considered slightly less important in group than in 1-to-1 conversation. This indicates that in the presence of a larger group, personal emotions might occasionally take a backseat, giving priority to other aspects of conversational success.

2.5.3 Strengths and limitations

The Group Concept Mapping (GCM) method offers flexibility, enabling each participant to engage in brainstorming, sorting, and rating activities individually or collaboratively in a group setting. In this study, participants worked individually rather than in a group. This approach had the advantage of preventing participants from being influenced by each other's ideas, though it is possible that collective brainstorming using the same

prompts might have generated a slightly different result. The current study managed to integrate viewpoints of older adults with normal and impaired hearing amid the challenges posed by the pandemic context of SARS-CoV-2. This was possible by conducting all three participatory steps through web-based activities. Another strength of this method lies in its capacity to minimize researcher influence, allowing participants to group ideas in a manner that best aligns with their individual perspectives.

The refinement and condensation of statements originating from the brainstorming phase is an essential step of the process for several reasons. Raw statements can contain information that would impede the sorting and rating activities. For example, raw statements might hold irrelevant answers to the focus question, personal information, jargon elements, typos, repetitive ideas, or multiple ideas within the same statement. Also, managing a large number of statements during sorting and rating activities might become a burden for participants. Rosas and Kane (2011) found that an average of 96.3 statements have been used in more than sixty GCM studies. In the present study, we began with a pool of 263 statements, which was subsequently condensed to 73 through trimming. While trimming is a necessary step, it comes with some drawbacks. For instance, recurring statements often highlight prevalent ideas within the sample group. During the trimming process, a frequently occurring idea may thus lose its prominence. To illustrate, within the initial statements of this study, we identified 16 statements pertaining to the need to request repetition (e.g., ‘Not having to ask someone to repeat what was said’). However, post-trimming, this number reduced to just 3. It is plausible that the prevalence of these ideas might be reflected in their importance ratings. However, if a concept is highly common but not of great importance, it raises concerns that valuable information may be lost in the process. Also, some statements were edited in the trimming process. For example, ‘the person who is talking has a loud and clear speaking voice’ was edited to ‘the speaker has a loud and clear speaking voice’. These edits led to several statements commencing with the phrase 'the speaker,' and this might be a reason why they have been grouped together by the participants.

One potential limitation of this study is the relatively small number of participants, as well as the homogeneity of the participant pool. It is conceivable that a more diverse and larger group could have generated a wider array of responses. However, the study deliberately focused on a specific demographic, namely individuals aged 50-75 residing in Glasgow, which was drawn from a relatively homogeneous participant pool. This targeted approach was chosen with the awareness that broadening the participant criteria, such as including younger adults, might have produced different priorities in conversation and potentially led to a distinct concept map for the same target concept. Moreover, it's worth considering the potential influence of participants' retirement status. The majority of participants (n=20) were retired individuals. Retirement can introduce limitations on the type, frequency, and settings of one's interactions. Therefore, it's possible that their perception of conversation success may be shaped by their retired status. It is also important to acknowledge that the normal hearing participants in this study were drawn from the same participant pool as the impaired hearing group. Individuals become part of this pool if they have undergone an audiological assessment at Glasgow Royal Infirmary within the past five years and have consented to be part of our participant panel. Consequently, while their audiograms indicated normal hearing from an audiometric perspective, it's possible that they may have had other hearing-related concerns or complaints.

The study, conducted from February to June 2021, coincided with various phases of global pandemic-related curfews. This temporal alignment could have shaped participants' responses, with face-to-face interpersonal communication being affected by social distancing measures. Given the constraints of isolation rules, participants may have found themselves engaging in fewer in-person conversations than they typically would. Consequently, heightened feelings of loneliness might have introduced a potential bias, influencing participants to evaluate aspects of conversation differently than their usual perspectives would suggest.

2.6 Conclusion

The present study aimed to determine how PwIH and PwNH understand conversation success in 1-to-1 and group conversations. Seven unique themes that relate to conversation success were identified: (1) *Being able to listen easily;* (2) *Being spoken to in a helpful way*; (3) *Being engaged and accepted*; (4) *Sharing information as desired*; (5) *Perceiving flowing and balanced interaction*, (6) *Feeling positive emotions*; (7) *Not having to engage coping mechanisms*. The second aim of this study was to explore how these themes differ in 1-to-1 versus group conversations. The results of this investigation show that *Being able to listen easily*, *Being spoken to in a helpful way* and *Sharing information as desired* are significantly more important in group than in 1-to-1 conversations for all participants, regardless of their hearing ability. No significant differences were found between PwNH and PwIH in terms of importance of these themes.

Our findings suggest that perceived conversation success is a multifaceted concept that encompasses elements of the classic communication chain (speaker, channel, listener) together with subjective components such as feelings and coping strategies. These results bring us closer to understanding whether and how conversation success can be quantified. This framework will serve as the foundation for forthcoming studies which will investigate how these elements apply in face-to-face group interactions.

Chapter 3

A group conversation experiment (Study 2): method

Portions of this chapter overlap with the following article submitted for publication:

Raluca Nicoras*, Bryony Buck*, Rosa-Linde Fischer, Matthew Godfrey, Lauren V. Hadley, Karolina Smeds & Graham Naylor, “Effective design for experiments on small-group conversation: insights from an example study” submitted to the Small Group Research journal.

*joint first authors

This overlap is primarily seen in the procedure section. However, most of the content in this chapter is distinct, offering a more in-depth view of the entire methodology including detailed descriptions of participants as well as the tools used throughout data collection.

The data collection for Study 2 was a collaborative effort involving Dr. Bryony Buck and Matthew Godfrey.

3.1 Chapter Summary

In this chapter I describe the methodology of the second study of this PhD thesis which was a laboratory experiment involving 18 quartets (groups of four people) composed of PwNH and PwIH. Overall, the study aimed to investigate the self-perception of conversation success in various background noise levels and aiding status conditions (analysis and results presented in Chapter 4) and link it to participants’ communication behaviour (analysis and results presented in Chapter 5). This chapter focuses only on the rationale and methodology behind the study's design, thereby providing context for the analyses presented in subsequent chapters. The data collection resulted in 108 video-recorded conversations along with audio, motion tracking data and continuous participant feedback.

To contextualize the experimental design within the broader topic of conversation success, I first outline the motivations behind the design choices. Then, I describe the design and the procedure, including challenges faced in

data collection. Finally, I report the feedback received from the participants during the debrief sessions and I focus on what can be learned from running such a complex experiment with groups of four participants in a laboratory setting.

3.2 Background

Previous laboratory measures of conversation have singled out different behaviours such as voice level, turn transitions, gestures, and head movement (Beechey et al., 2020; Hadley et al., 2019, 2021b; Holler & Wilkin, 2011). While specific behaviours have been studied in relation to background noise changes, hearing aid use, and hearing status, there is no clear understanding of how these behaviours relate to conversation success.

One aspect of conversation behaviour that could relate to conversation success is the ability to share information successfully. Paradigms that tackle this aspect typically make use of storytelling or task-oriented conversations. The performance is typically measured by how accurately a listener understands a story they have been told or the time taken to complete game-based tasks like Diapix/Map/Maze or Tangram puzzles. Game-based methods, in comparison to storytelling, offer a more interactive experience, demanding a level of participant engagement. This methodology has been proven effective in experiments involving healthy individuals (Stephens et al., 2010), and also with specific groups such as those with aphasia (Ramsberger and Rende 2002) or hearing loss (Beechey 2018). Importantly however, while the approach is well-suited for laboratory settings, it might restrict the spontaneity found in everyday conversations.

Apart from efficient information exchange and performance in task-oriented games, speaker-listener alignment has also been interpreted as an indicator of conversation success. Similar behaviour between interlocutors could indicate similar representations and common ground therefore a better understanding. Alignment indicates parity between participants and has been studied at different levels including neural – e.g., when the speaker's and listener's neural activity is synchronised (Anders et al., 2011; Stephens et al., 2010), linguistic – e.g., when the listener is repeating the words of the speaker

(Garrod & Pickering, 2004), and behavioural – e.g., when the listener imitates the speaker's gestures (Holler & Wilkin, 2011).

Coordination of turns between speakers can also be an indicator of the quality of a conversation. Lack of coordination can result in overlap in turns which occurs when participants are talking over each other. One of the bases for coordination is the listener's ability to predict when it is their turn to speak, in essence, the ability to anticipate when the current speaker will stop talking. It has been shown that people are good at predicting the end-of-turn of a speaker in general (De Ruiter et al., 2006) and are even better at predicting the end-of-turn of speakers whose production style is similar to their own (Hadley et al., 2020). The method requires a participant to listen to a recorded statement and promptly press a button when they anticipate that the statement ends. Participants are given explicit instructions not to wait until the sentence is completed but to press the button as soon as they think that the sentence ends. Similar to information exchange tasks, this method is easy to reproduce in the laboratory settings and provides valuable evidence on how participants manage to coordinate. However, natural conversations are complex and require listening to more sentences in a row, following a narrative, preparing to respond, and detect multimodal backchannels, all while coping with potential environmental adversities.

Other studies explored the idea of fluency as a measure of conversation quality. Some factors that intuitively could be expected to compromise conversation fluency are long pauses, miscommunication, and the need for repairs. Pauses, in this context, refer to periods in a conversation without vocal activity that exceed the typical gaps between speech turns, which are usually brief (on average 200ms in English) (Stivers et al., 2009). Existing literature supports the idea that long pauses make people feel uncomfortable except in cases where the conversation takes place in a therapeutic context or among familiar interlocutors (Templeton et al., 2023). On top of long pauses, miscommunication and repair can also perturb the conversational flow. Miscommunication occurs when, during a conversation, at least one participant detects a failure in the transmission or comprehension of information. Repair,

in this context, refers to all the actions taken to rectify or address the breakdown in communication (Lind, 2010). Previous work on conversation has analysed how often breakdown and repair take place, especially in patients with hearing-loss (Lind et al., 2006). Breakdown and repair seem to have detrimental effects on fluency, conversations involving fewer breakdowns being considered more desirable. For instance, children with cochlear implants are perceived more favourably when they spend less times in communication breakdowns (Tye-Murray, 2003). In situations with conflicting information, a repair isn't always required (Stalnaker 2002). If mutual understanding isn't affected, the absence of a repair isn't problematic (e.g.: in a bustling pub, Marry and Bill are placing their drink orders. Marry takes the lead, asking Bill for their drink preference. Bill requests sparkling water. However, Marry immediately requests a bottle of still water from the bartender. While Bill recognizes the misunderstanding, they decide that initiating a repair is not necessary in this situation, and they decide to avoid the repair.). Furthermore, even when a repair is necessary, some argue that well-managed miscommunication does not necessarily hinder conversation fluency. Instead, it can foster interactional flexibility (Healey, De Ruiter, and Mills 2018).

While all these studies provide insights into how to measure different aspects of a conversation, two major aspects are overlooked: what do the interlocutors themselves think and what behaviours are associated with their perception?

In an earlier study (Chapter 2), I explored the notion of conversation success from the perspectives of PwNH and PwIH. The investigation yielded seven distinct factors that drive conversation success (*Being able to listen easily, Being spoken to in a helpful way, Being engaged and accepted, Sharing information as desired, Perceiving flowing and balanced interaction, Feeling positive emotions*, and *Not having to engage coping mechanisms*). However, these represent hypothetical ideas created by participants during an online study, responding from the comfort of their homes which can be different than their perception during a real conversation. The task itself was simple. Participants were prompted to answer the following questions: ‘What does

"successful conversation" look like?' and 'Think about a successful conversation you have taken part in. What aspects of that conversation contributed to its success?' However, this accessed participants' thinking processes, memory, and imagination, rather than their perception during an active participation. In daily life, conversations happen "here and now" and require spontaneous reactions. An online-only approach to studying the concept of conversation success lacks critical elements such as authenticity, dynamism, and expectation-reality conflict experience. Participants' thoughts and self-perception might differ when they are passively assessing the success of a past conversation in contrast to rating the success of a conversation they have in-the-moment.

In-the-moment assessment can be successfully undertaken through various approaches including laboratory experiments or ecological momentary assessment (EMA) methods. Assessing moment-to-moment perceptions in laboratory experiments offer several advantages, including precise control over environmental variables and stimuli, ensuring replicability, and minimizing the influence of confounding variables. This control allows us to establish causal relationships between variables that could influence success with a greater confidence. Additionally, laboratory settings provide access to necessary equipment and resources, facilitating standardized data collection procedures. However, it is essential to acknowledge the limitations of laboratory studies. The artificial context of the laboratory may compromise ecological validity (Keidser et al., 2020), as participants' behaviours and responses may differ in real-world settings. Participants' awareness of being observed or the experimental setup itself may influence their behaviour, potentially introducing biases into the data. Despite these limitations, laboratory-based approaches remain valuable for their ability to maintain experimental control and facilitate rigorous hypothesis testing. Thus, we have selected a laboratory-based experimental approach as the optimal method to achieve our research objectives detailed below. Through this approach, we aim to advance the investigation of conversation success by examining in-person face-to-face live interactions.

This study was based on the premise that different levels of conversation success could be achieved by manipulating conversations to alter how interlocutors perceive its success, which in turn can trigger associated behaviours. The variables were chosen based on the results from Study 1 (Chapter 2). The most important conversation success theme was *Being able to listen easily* which contained statements like 'Hearing aids working (e.g: not whistling, coming loose or batteries going flat).' and 'Being in a quiet place without background noise or music.'. These findings suggest that wearing or not hearing aids during conversations in different background noise levels could alter the perception of conversation success.

Extensive research has demonstrated that background noise can interfere with speech perception and understanding (McArdle & Wilson, 2009), making conversation more difficult, increasing fatigue (Holman et al., 2019), communication effort (Beechey et al., 2020) and eliciting behavioural changes such as Lombard speech (Bottalico et al., 2022), leaning in as well as relying on lip reading (Hadley et al., 2021). In this study, three background noise levels were used: low (30 dBA, with loudspeakers muted and only in-room equipment noise present), medium (54 dBA babble), and high (72 dBA babble). The low noise level at 30 dBA, would be inaudible for most listeners with at least moderate hearing loss and would likely lead to conversations perceived as successful. The medium noise level at 54 dBA resembles a populated yet not overly noisy social environment, and thus would be expected to lead to a lower conversation success compared to the low noise level. The high noise level at 72 dBA corresponds to the noise level typically found in a medium-occupancy restaurant (To & Chung, 2015), and conversations in this setting would be anticipated to be rated as the least successful. The analyses presented in Chapter 4 (RQ3) will reveal if our expectations were met.

The communication difficulties can increase when interlocutors have hearing loss (Kiessling et al., 2003). However, when provided with amplification through hearing aids, PwIH seem to behave more similarly to their normal hearing counterparts (Petersen et al., 2022), potentially perceiving conversation success more positively when aided. Therefore, in this study, we

contrasted the conversation success perceptions and behaviors of PwIH and PwNH in an experiment where PwIH wore their hearing aids for half of the time.

Conversation characteristics such as size, familiarity, and goals could impact how interlocutors perceive conversation success (Chapter 1). In this study, these factors were deliberately held constant. The experiment exclusively involved four-way interactions among unacquainted interlocutors engaging in natural conversations. Below, I motivate our choices.

Compared to group conversations, dyads (one-to-one) tend to be more intimate, with one person holding their conversation partner's undivided attention. Whereas in conversations with three or more people, there are multiple speakers to keep track of and engage with. This results in increased head movements and gaze variability, less predictable speaker sequences, and more intense competition for floor-taking (Cooney et al., 2020; Hadley et al., 2021b). The unpredictability of these interactions, characterized by rapid speaker or topic changes, exacerbates challenges for those with hearing impairments (Vas et al., 2017). Following the conversation can prove demanding, especially when compounded by overlapping speech and heightened acoustic interference. On the other hand, multi-party interactions also have some advantages. When there are more than two interlocutors, it can relieve the pressure of always being the next speaker, making it easier for individuals to employ withdrawing strategies. Simultaneously, there is room for an alliance between interlocutors, and repairs can be solved by other listeners (observers) not only speakers. Although research shows that group conversations occur less often than dyadic conversations (Paperkoorn, 2020), they are considered very important by hearing-aid users (Wolters et al., 2016). So that is why I focus only on group conversations in the current and all following chapters. Moving from dyads to groups, experiments become disproportionately more complex to design and execute, including the demands on technology, researchers, space, time, and logistics. As group size increases, the number of possible interactive effects inflates drastically. Furthermore participants may perceive greater social pressure to conform (Bond, 2005),

posing a potential ethical issue, as participants may be reluctant to withdraw consent or request a break. All these challenges highlight the importance of thorough design, planning and execution for successful experiments on group conversations. In this study, we concentrated on quartets (groups of four people) which aligns with prior research indicating that group conversations commonly comprise four individuals (Krems & Wilkes, 2019). Additionally, interactions involving four participants can fragment, resulting in the possibility of two simultaneous conversations occurring. By including two individuals with normal hearing and two individuals with hearing loss in the group of four, we can observe whether fragmentation occurs between individuals of the same or different hearing status.

Familiarity among interlocutors could alter their perception of conversation success. While familiar interlocutors may share past experiences and feelings, conversations between strangers can be surprising due to the lack of expectations. To remove any emotional bias associated with shared past experiences, this study focused exclusively on unacquainted sets of participants. This choice also facilitated the recruitment process, as finding people with desired similar characteristics might have been challenging.

Literature differentiates between conversations that are informational and have a concrete goal, and relational conversations with a social goal such as bonding (Clark et al., 2019; Yeomans et al., 2022). Creating conversations that are purely instrumental or purely social can be unrealistic, as natural conversations typically combine both aspects. Therefore, the topics used in this study were deliberately developed to encourage light-hearted engagement rather than serious, goal-oriented (potentially competitive) performance. We expected the resulting conversations to exhibit a mix of transactional and relational characteristics (Yeomans et al., 2022).

With group size, familiarity among interlocutors, and conversation goals held constant, the experimental contrasts in this study were informed by participant responses in Study 1. These contrasts involved three independent variables: hearing ability, background noise level, and hearing aid usage. This manipulation aimed to induce a substantial shift in participants' perception of

their conversation success (Chapter 4) and enable us to link participants' perceptions to their behaviors (Chapter 5).

3.3 Aims

The overarching goals of the current study were to understand 1) how people perceive success in conversations and 2) how their behaviours relate to their perception of conversation success. This dataset enables the categorization of conversations along a continuum, ranging from successful to unsuccessful conversations.

3.4 Method

This section outlines the research design employed to meet the previously mentioned aims, encompassing participant recruitment, experimental design, and procedural aspects.

3.4.1 Participant Recruitment

72 participants (age mean = 67.86, SD = 7.5, 39 females) completed the entire study out of a total of 126 participants who completed only the first session. All participants were recruited from the Hearing Sciences Scottish Section (HSSS) participant pool under the following criteria:

- Age range: 50 to 75 years.
- Fluent in English.
- Self-report of normal or corrected to normal visual acuity.
- No medical condition that could impact their communication abilities (beyond hearing impairments).
- Left-right asymmetry of four-frequency (0.5, 1, 2, 4 kHz) pure-tone average hearing threshold (4F-PTA) below10 dB HL.

Additional inclusion criteria specifically for PwNH:

- Worse-ear 4F-PTA below 20 dB HL

Additional inclusion criteria specifically for PwIH:

- Better-ear 4F-PTA between 41 and 65 dB HL.
- Bilateral hearing aid users.

- Experienced hearing aid users, defined as individuals who had been wearing their hearing aids for more than 6 months and had self-reported wearing them for at least 2 hours daily.

Recruitment and scheduling difficulties forced us to include a few participants who fell slightly outside these ranges on one or more criteria. This resulted in the age range being expanded to 50 – 78 years, the worse ear 4F-PTA limit for PwNH extended to 28.75 dB HL, the better ear 4F-PTA range for PwIH extended to 35 – 65 dB HL, and the 4F-PTA asymmetry range expanded to 15 dB HL. Furthermore, one participant had only three months experience of HA use.

The experiment unfolded in two sessions. Participants were paid £10 per session for taking part. This research received ethical approval from the West of Scotland Research Ethics Committee (18/WS/0007) and the NHS Greater Glasgow and Clyde R&D (GN18EN094). Recruitment during the COVID-19 pandemic brought additional challenges. For instance, individuals who expressed serious concerns about the potential transmission of the virus were not invited to the second session. Additionally, to confirm their attendance and assess any symptoms related to COVID-19, participants were contacted before each session.

3.4.2 Experimental design

Manipulated Variables: Hearing ability as a between-subject factor divided participants into two groups based on their hearing status: PwNH and PwIH. Background noise levels were manipulated as a within-subject factor at three levels: low (30 dBA), medium (54 dBA), and high (72 dBA). Furthermore, the use of hearing aids was examined as an additional within-subject factor, with PwIH using their hearing aids in half of the conversations. This action was required only from participants in the impaired hearing group; however, it was considered to impact all participants involved. As a normal hearing participant, conversing with PwIH while they wear their hearing aids might differ from when conversing with the same people while they are unaided. Therefore, these experimental conditions are referred to as “unaided” even though they include both PwNH and PwIH. The order was

counterbalanced for both hearing aid use and the background noise level in order to avoid order effects.

This results in a 2x3x2 mixed-factorial design. Further information regarding the background noise's characteristics, duration, and source will be provided in the technical execution section.

Measured Variables: The study employed various measurements to capture different aspects of conversation perception and behaviours, with the aim of linking participants' self-perception of conversation success to their communication behaviour under different conditions. This led to two categories of measurements:

a) Self-perception: Participants provided continuous feedback on their perception of conversation success through interface tablets, addressing this first specific aim regarding perception of conversation success. Additionally, post-conversation feedback was collected via an 11-item survey displayed on the same tablets.

b) Objective behaviours: Audio recordings were obtained using individual microphones to capture vocal activity, allowing for the calculation of speech levels, speech proportion, and detection of laughter. Video recordings were acquired using multiple cameras to capture facial expressions, eye-gaze, and gestures. Motion capture systems recorded head movement data, enabling measurements such as head orientation and proximity. These measures addressed the second specific aim regarding behaviours relating to success.

Further technical details about the data collection are provided in the procedure section.

3.4.3 Procedure

In this section, I describe the experiment execution providing chronological details and information on technical setup. Prior to the experiment execution we piloted the study and conducted a patient and public involvement (PPI) meeting.

Several changes were made based on the feedback received. The duration of each experimental condition (conversation) was reduced from 8 to

6 minutes, as participants found the initial length overly prolonged. Additionally, the tablet position on the table rather than on the chair was considered optimal. The PPI members provided feedback on the participant information sheet (PIS) and experimental instructions. Also, the comfort level of the experimental setup for participants was addressed, such as adjusting seats, lighting, temperature. The PPI meeting allowed us to shape the experimental set-up to align with the needs of participants.

3.4.3.1 Session one

The first session was done with single individuals, and included signing the consent form, completing the hearing assessment and a set of questionnaires, and being familiarised with the second session set-up (Table 3.1). These are described in detail below.

Table 3.1 Details of the first and second sessions

Session	First session	Second session
Duration	1h	1h30 - 2h
Type	Individual	Group
Chronological events	Hearing test	Calibration and instructions
	Questionnaires	Training
	Familiarisation session	Experiment*

*Experiment details in Figure 3.1

3.4.3.1.1 Hearing assessment

The hearing assessment consisted of a standard audiogram performed in a sound-attenuated booth to verify participants' hearing thresholds, which might have changed since their previous hearing test. The hearing assessment continued with two self-reported measurements: Hearing Handicap Inventory for Adults/Elder (HHIA/E) (Ventry & Weinstein, 1982) and Speech, Spatial and Qualities of Hearing scale (SSQ12) (Noble et al., 2013). While a higher score on the HHIA/E indicates a higher perceived handicap due to hearing loss (i.e., the individual perceives their hearing loss to be more of a handicap in

their daily life), the SSQ12 scales such that a higher score suggests better hearing ability or less perceived impact from hearing loss. Participants entering the HSSS participant database have previously been to an audiological check at Glasgow Royal Infirmary and have given their consent to be part of our research panel. Hence, while their audiograms can show normal hearing thresholds, they may nonetheless have had other hearing-related complaints. The two questionnaires mentioned above aimed to help us detect these participants and avoid including participants with high perceived hearing handicap in the normal hearing experimental group.

3.4.3.1.2 Covariates and questionnaires

In multi-party interaction studies, the number of covariates directly increases with the diversity of individual characteristics involved. Usually, by controlling covariates researchers attempt to isolate the effects of the independent variable on the dependent variable, reducing the risk of confounding results. Since it would be impossible to account for all potential covariates, a compromise is needed. In this experiment, participants' self-perception and behavioural patterns during conversations hold crucial importance. Factors such as affective mood, autistic traits, feelings of loneliness, and empathy could all potentially impact individuals' perception of conversation success and their corresponding behaviour. Consequently, the following set of questionnaires was administrated to ensure that IH and NH groups were similar in positive/negative affect, autistic traits, empathy and loneliness.

- Positive and Negative Affect Scale (PANAS) is a 20-item questionnaire aiming to measure positive and negative emotions. Research suggests that positive/negative affect can impact perceptions of interaction quality (Berry & Hansen, 1996).
- Autism Spectrum Quotient Short version (AQ-10 adult) is a 10-item questionnaire used to measure autistic traits in adults. While the AQ-10 doesn't confirm a diagnosis of autism, it indicates that a person shares traits with individuals on the

autism spectrum. In interactions, individuals with high AQ scores could show potential social communication deficits and they could display atypical empathic responses (Harmsen, 2019).

- Interpersonal Reactivity Index (IRI) is a 28-item questionnaire aiming to measure empathy, through four subscales. In this study we assessed the following subscales: perspective-taking, empathic concerns, and fantasy. This choice was informed by the idea that perspective-taking can enhance social coordination and behavioural mimicry, both potential key elements in successful conversations (Galinsky et al., 2005). The personal distress subscale was excluded from this study as it measures self-oriented feelings of personal anxiety, which lie outside the scope of the study.
- Loneliness Scale (UCLA) is a 20-item questionnaire used to measure feelings of loneliness and social isolation. This questionnaire is relevant for interaction studies as it has been noted that individuals with higher scores can be evaluated by others as being less conversationally competent (Spitzberg & Canary, 1985).

3.4.3.1.3 Familiarisation session

Ten minutes were dedicated to familiarisation with the setting in which the second session would take place. Participants were welcomed into the experimental room and then presented the equipment. They were given a tablet with the Emotouch software (Louven et al., 2022) showing a slider going from "not at all successful" on the left side of the screen to "very successful" on the right side. They could use a finger to drag their position on the slider or they could use a finger to tap on different places of the slider.

3.4.3.2 Session two

For the second session, 72[1] participants were invited in as 18 mixed-gender and mixed-hearing ability quartets (two males and two females[2] / two normal and two impaired hearing participants[3]). Participants were unfamiliar with each other before the experiment. Each quartet engaged in a 2-minute training conversation in quiet, followed by a series of six conversations of six minutes each. Background noise was presented at three different levels: low (30 dBA), medium (54 dBA), and high (72 dBA). These were split over two equal blocks. In one block, IH participants used their HAs, in the other they did not (order was counterbalanced across quartets). Each block contained all three noise conditions, with the order counterbalanced across quartets and aiding conditions. The second session took approximately 1h30 with a comfort break between the two blocks (Figure 3.1).

[1] In total, 76 participants (19 quartets) completed session two, but one quartet was removed due to lost data (see table 3.3)

[2] Due to recruitment difficulties, 5 quartets had a 3:1 gender distribution instead of 2:2.

[3] Two quartets required the substitution of one participant with an HSSS staff member matching in age. While one matched the hearing threshold requirements, the other simulated the hearing loss, in the "unaided" condition, by wearing earplugs.

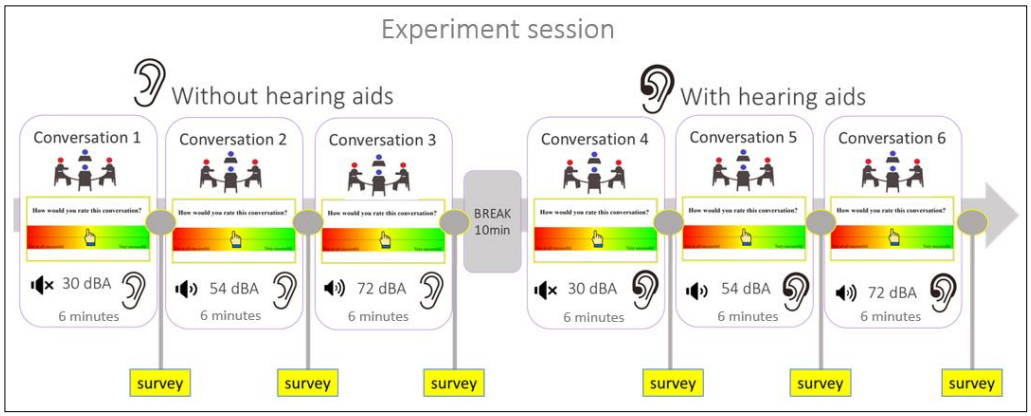


Figure 3.1 Description of the second session: six conversations corresponding to six conditions; each quartet includes two PwIH (red) and two PwNH (blue). This represents an example. The order of the conditions was counterbalanced in the experiment.

3.4.3.2.1 Room considerations

The experiment took place in a sound-attenuated room (4.3 x 4.6 x 2.6 m), with a carpeted floor, and walls and ceiling covered with sound-absorbing foam wedges. The room had only one door and one window. The four participants were seated around a table (diameter 1m) in the centre of a square of 4 equidistantly spaced loudspeakers (Tannoy VX-6) (figure 3.2). Four other phantom loudspeakers were in the room, but did not produce any sound and were only used as camera holders. In order to prevent participants leaning to one side for comfort reasons (rather than to hear better), chairs without armrests were used. Opposite chairs were 2 m apart and adjacent chairs 1.4 m apart. The distance between chairs was checked and readjusted if needed at the beginning of each conversation.

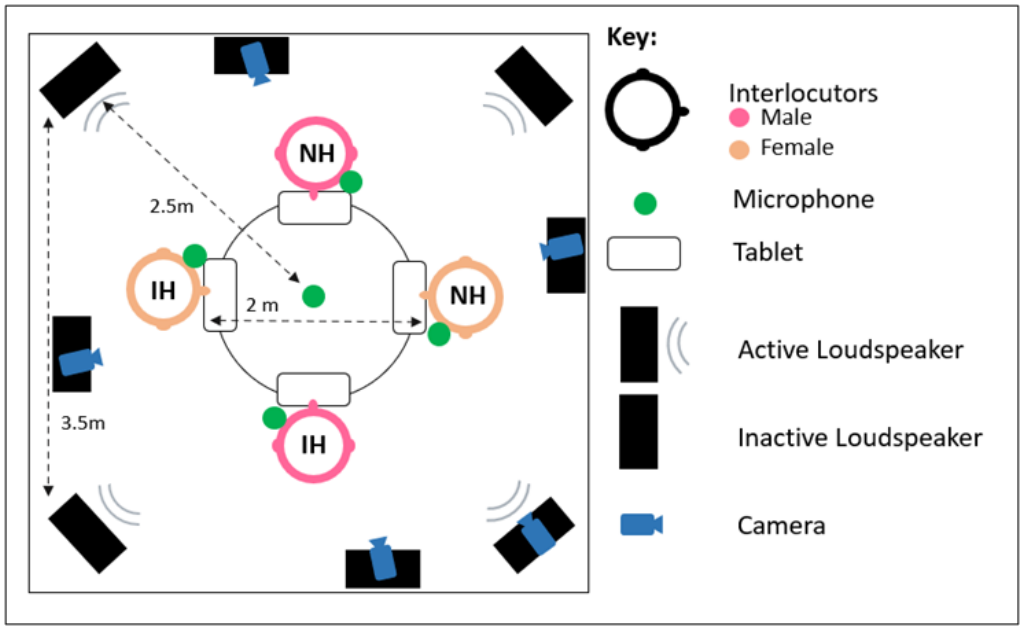


Figure 3.2 Schematic plan of the experimental setup.

3.4.3.2.2 Quartet considerations

Our experimental design included unacquainted sets of participants. This choice comes with the potential consequences of heightened anxiety levels among participants and the possibility of order effects, as participants may become more relaxed in each other's company over time. Conducting the study during the global COVID-19 pandemic presented additional challenges for running multi-party studies in-person after two years of consecutive worldwide lockdowns. The physical distancing between participants was determined based on the regulations in place at the time in Glasgow, Scotland. Another important aspect was the decision to not allow face masks during the experiment, which was communicated in the participant information sheet and the first experimental session. However, this choice carried the risk that some participants decided to withdraw from the study after the first session.

All quartets in our study were composed of two PwIH and two PwNH. This allows us to explore mutual influences in their behaviours and, at the same time, avoiding imbalance. For instance, while a quartet formed by three PwNH and one PwIH is realistic, it might trigger unpleasant feelings for the latter. On

the other hand, a quartet formed by one PwNH and three PwIH is less common and not a priority for the current study.

When matching participants into quartets, several aspects were considered: gender, age gap and audiogram dissimilarity. We prioritised gender and age gap over audiogram differences. This was possible due to the strict hearing ability criteria mentioned above. However, recruitment flexibility was required to a certain extent in order to obtain the desired number of participants.

Regarding the physical arrangement, two configurations are possible for pairing 2 PwNH and 2 PwIH in a circle: either matching hearing abilities and differing genders facing each other, or matching genders and differing hearing abilities facing each other. We opted for the latter configuration to allow all participants to engage in sub-conversations with neighbours of either hearing status.

3.4.3.2.3 Conversation topic

The participants' task was to converse starting from a scenario shaped such as it required a consensus answer without requiring specific prior knowledge (e.g.: "Imagine that life on land is no longer possible. In order to survive you either have to live underwater or in space. Discuss which you would prefer as a group."). The full list of topics used in this experiment is attached in Appendix C. Our goal was to ensure that all participants had an equal opportunity to contribute to the conversation.

In order to ensure that all participants had equal chances to contribute to the conversation, the given topic of conversation that does not require specific prior knowledge was needed. For instance, in paradigms such as "timeline ranking of events" participants need to access their knowledge on the chronology of certain events to be able to engage in conversation. In contrast, hypothetical consensus topics require no specific expertise and factual knowledge will not close the conversation. While it is possible to use facts in order to say, for instance, that it is better to live underwater than in space, the topic accesses the personal preferences and participants' imagination which

makes possible the continuation of the discussion around the given topic beyond the factual knowledge. Also, an optimal conversational starter should avoid any potential invested controversy (e.g.: religion, politics, etc.). While we considered the consensus topic the optimal task for our experiment, it also brings some drawbacks. They present hypothetical situations that may not be meaningful for participants and might present topics that participants would not talk about in their daily life. This increases the risk of low ecological validity.

3.4.3.2.4 Background noise

The same background noise was presented continuously for 6 minutes corresponding to each conversation. In order to attenuate any potential surprise and to give hearing aids enough time to accommodate, smoothing was applied at the beginning and the end of each conversation for 10 seconds. Previously used by Hadley et al. (2019, 2020), the eight-talker (four females, four males, all with British accents) babble noise was generated by concatenating sentences from Stacey and Summerfield (2007). In the current experiment, eight talkers were combined into 4 groups, one group corresponding to one loudspeaker. Two female voices overlapped with two male voices per loudspeaker. In restaurants people at the tables around can be localised closer and further away. That is why, 2 dB difference was added between two sets of pairs of each loudspeaker. A Bruel & Kjaer sound level meter was used to ensure that the played sound matched the desired level for each condition.

3.4.3.2.5 Measurements

In order to address both perceived success (Chapter 4), and behaviours related to that success (Chapter 5), we recorded both self-reported perception of conversation success as well as communication behaviours.

3.4.3.2.5.1 Subjective perception

Continuous feedback during each conversation

Continuous feedback during conversations was collected using a slider ranging from “unsuccessful” at one extreme to “very successful” at the other

extreme. Participants were instructed to rate how successful they perceived the conversation during the conversation, and they were reminded between the conversations to interact with the slider as often as possible. This offers a collection of data points from which we can infer instances that triggered a change in subjective perception of conversation success.

Feedback after the conversation

The post conversation survey asked participants for an overall view of the success of the conversation they just finished. This makes possible the comparison of their perception during the conversation versus the post-factum evaluation. Additionally, this assessment provided information that was not possible to be collected during the conversation. Participants were given an 11-item survey after each conversation. The items in this survey were chosen based on the seven themes of conversation success (Chapter 2). Each was reflected through one or more survey items (see Table 3.2). While all items measured an aspect that was previously shown to drive conversation success, one additional item specifically asking about active behaviours to make conversations successful was added (item 2, see table 3.2).

Table 3.2 The survey administered after each condition.

Motivation	Survey item	Rating scale
Overall conversation success	1) To what extent would you rate this conversation as successful?	1 = unsuccessful 5 = successful
Communication effort (Beechey, 2020)	2) Did you do anything or say anything in a particular way to make this conversation successful?	1 = not at all 5 = very much
Seven facets that drive conversation success (Chapter 2)		
Being able to listen easily	3) How easy was it for you to follow the conversation?	1 = not easy 5 = very easy
Being spoken to in a helpful way	4) To what extent was the person sitting on your left speaking in a helpful way? 5) To what extent was the person sitting in front of you speaking in a helpful way? 6) To what extent was the person sitting on your right speaking in a helpful way?	1 = not helpful 5 = very helpful
Being engaged and accepted	7) How connected did you feel with other participants?	1 = not connected 5 = very connected

Sharing information as desired	8) To what extent did you share information successfully?	1 = not successfully 5 = very successfully
Feeling positive emotions	9) How enjoyable was this conversation for you?	1 = not enjoyable 5 = very enjoyable
Perceiving flowing and balanced interaction	10) To what extent was this conversation flowing smoothly?	1 = not smoothly 5 = very smoothly
Not having to engage coping mechanisms	11) How often did you feel uncomfortable/anxious during this conversation?	1 = never 5 = always

Tools

Subjective ratings of conversation success were recorded using four 12" Microsoft Surface Tablets positioned on the table in front of each participant. EmoTouch software (Louven et al., 2022) was used to synchronise the screen displayed on all tablets and record participants' answers. During each six minute conversation, the question "How would you rate this conversation?" was displayed on the screen. Participants responded by moving the slider between left (0 = 'not at all successful") and right (50 = "very successful") (Figure 3.3). The scale points were invisible to the participants and were chosen such that the slider would have enough data points for the pointer to move smoothly on the screen, and also to offer a range that would be comparable to the data gathered from item 1 of the survey. Between conversations, the same software was used to prompt participants to complete the survey. For this part, the interface was programmed in self-paced mode, such that each participant could advance through the survey independently from others.

Data linked to the participants' subjective perception corresponds to the first aim. The analysis and results will be presented in Chapter 4.

3.4.3.2.5.2 Communication behaviour

The behaviours people show during conversations can be displayed simultaneously (e.g.: raising eyebrows while talking), sequentially (e.g.: talking and then laughing), in synchrony with their conversational partners (e.g.: two people nodding at the same time), or complementary with their

conversational partners (e.g.: one person talking, the other person laughing). In order to measure these multimodal behaviours, this study focuses on capturing cues that are visible and audible for an external observer: audio (speech), head motion and video (facial expressions and gestures). Collecting data from multiple sources requires careful calibration and synchronisation.

Speech

The speech was recorded using 4 microphones (DPA CORE 4288 Flex Headset Microphone) placed close (2-3cm) to each participant's mouth, and a microphone placed in the middle of the table. The audio was recorded at 44.1 kHz. In order to monitor the experiment from outside, an additional microphone was placed in the room such that the audio was streamed in real-time to the researchers' room. This microphone was controlled using a Raspberry Pi. After each participant microphone was attached and placed correctly (at the beginning of the 1st conversation and after the break), participants were asked to read out loud a phonetically balanced statement, known as a Harvard sentence (from IEEE Recommended Practice for Speech Quality Measurements, 1969) (e.g.: "The source of the huge river is the clear spring."). Recording participants speaking these sentences without background noise serves as a reference for their speech level during the experimental conditions.

Head motion

The head position was recorded using the Vicon system, a marker-based motion capture system that employs infrared cameras to trace the three-dimensional positions of reflective markers strategically positioned on the body. Each participant wore a hat containing five detectable markers (referred to as a crown), in a room with eight Vicon Bonita B-10 infrared cameras. Using Vicon Tracker Software (version 3.9) the cameras recognised the markers positioned on each person's head. Head position was calibrated for each participant at the centre of their head and in relation to the centre of the room. The temporal sampling rate was 100 Hz and spatial resolution was under

0.01°. Vicon motion-tracking system is extensively used to track body movements in various domains such as sports (Monnet et al., 2014) and physical rehabilitation (Alarcón-Aldana et al., 2020). In hearing sciences, the system was shown effective in tracking head movement in listening tasks (Brimijoin & Akeroyd, 2012), and in dyadic and triadic communication (Hadley et al., 2021).

Video

Video data was recorded using four ELP 5-50mm zoom cameras oriented towards each person's face and one ELP Web Camera USB 5MP 2.8-12mm Varifocal Wide-Angle Lens capturing the conversation from one side of the room (figure 3.3). The video stream from all five cameras (figure 3.3) was recorded simultaneously using OBS Studio software at 1920x1080p and 30 FPS.

Software

MATLAB was used to control the background noise and record the speech and head motion data.

Figure 3.3 Experimental set-up with details of the tablet screen as seen by the participants during a conversation.

3.4.3.3 Debrief

The debriefing stage was structured in a twofold manner, with an immediate group debrief and a later individual follow-up either via phone or email. Neither of these components were formally part of the experiment and their participation was entirely voluntary.

The group debrief (N=19 quartets) was held immediately following the second session in the same location. The purposes of the study were clearly outlined to the participants, and an open floor for questions and feedback was provided. Participants were guided by a predetermined series of inquiries but were also encouraged to voice any additional thoughts. This format carried the benefit of capturing immediate reflections while participants' experiences were still vivid, facilitating an understanding of shared experiences among group members.

For the individual debrief (N=21), each participant was contacted through their preferred mode of communication—phone or email—several days post-experiment. This format offered a private space to express thoughts, appealing to those who preferred to divulge their experiences in solitude or who formulated insights over time. The debriefing process was guided by a set of questions such as “What was your overall experience of participation in this study?”, “How did you feel about the conversations you had with the other participants?”, “Would you be willing to take part in a similar study in the future, if we asked?”, “Do you have any other comments or feedback?”, but participants were invited to expand beyond these questions if they wished.

While the insights derived from the debriefing sessions are not incorporated into the statistical analysis, they offer a crucial perspective on participants' experiences, which influences the interpretation of our statistical outcomes. Therefore, these insights are succinctly presented in the Results section.

3.5 Results

3.5.1 Participants experience

Some concerns when running conversation research related to the difficulty of recording naturalistic behaviour, the ecological validity of the situations chosen, and the pressure participants might experience. Ideally, the task, the environment and the equipment should possess sufficient intrinsic engagement so that the participants behave as if they were not participating in an experiment, the situations presented in the experiment should be representative of the participants' real-life experiences, participants should not feel pressured to act in a manner that pleases the experimenter and should feel comfortable in their interactions with others, without the pressure to conform or reveal personal information and without feeling excessively anxious.

Determining whether this study successfully created a realistic experience for the participants can be challenging. To tackle this, we conducted debriefing sessions with participants to better understand their perceptions. The following provides an overview of some key feedback received.

Participants indicated that the experiment mirrored their daily experiences, with comments such as, "*I felt it really worthwhile as it highlighted the various situations in which I struggle to hear*", showing that our design captured a sense of realism enhancing the ecological validity of the study as defined by Keidser et al. (2020) "…the degree to which research findings reflect real-life hearing-related function, activity, or participation" . They perceived it as an enriching experience in spite of the difficulty, exemplified by feedback such as, "*I have been taking part in hearing science studies for around 15 years and found this experience really enjoyable and one of the best to-date...The atmosphere and background noise really did identify the issues that those with hearing difficulties suffer in day-to-day life*". In terms of equipment, the unanimous consensus among participants was that it was comfortable. Although some participants were initially aware of its presence, their focus on the conversations led them to forget they were wearing it.

Regarding the conversation topics, there was an overall agreement among participants that they were not appropriately suited for a middle-aged

demographic and therefore may not have captured typical conversation behaviour. One participant articulated: "*The questions we were asked to discuss weren't very suitable for middle-aged folks! (I think!) I think they'd suit younger people. I know you didn't mind what we talked about but it just meant it was harder to engage in conversation and even to want to engage in conversation*. ". Some participants expressed their preference for topics that are more realistic, similar to the one used in the training session ("Discuss how Glasgow has changed in the past 10 years."). While hypothetical scenarios can be entertaining, certain participants noted that these were subjects they would not typically discuss outside the experimental setting.

Although we aimed to devise topics steering clear of contentious subjects (e.g., religion, politics, COVID-19), avoiding controversy entirely proved unachievable in some conversations, exemplified by this comment, "*Enlightening - all strangers we agreed Scotland should be independent! That came out of the question: earth is unliveable - would you prefer to live in outer space or under the ocean?*". While certain individuals might appreciate or find intrigue in these types of topics, some participants reported discomfort, such as one who shared "*My only unease is where the chat turns to politics. I don't like to become involved in voicing my opinions in that situation.".*

Two participants described the experience as unpleasant. Despite the familiarisation period in the first session, the actual experience was reportedly more negative than anticipated. While they were told they can end the experiment at any time without giving a reason, they chose not to. One participant shared, "*The first conversation was especially hard - not knowing anyone, not being able to keep up, feeling a bit out of my depth - and I felt like I could dissolve into tears at the end of it. I have felt like this in real-life situations too.*" Yet, such feedback highlights that the experiment successfully emulated realistic scenarios in a laboratory setting.

3.5.2 Challenges and lost data

Compared to individual or dyadic studies, the orchestration of multi-party interactions presents an elevated demand for technical, human, and spatiotemporal resources. On the technical front, this involves not only an

increase in the amount of equipment but also ensuring synchronized, in-the-moment multimodal measurements. Coordinating participant schedules, minimizing the risk of no-shows, and handling on-site logistics during sessions necessitate detailed planning and effective coordination. These logistical intricacies further involve managing the timeline of participants' arrivals, their interactions with the researchers, and their involvement in shared spaces during calibration routines, all requiring the presence and support of multiple staff members.

Conducting an experiment that involves the simultaneous recording of multiple data streams, coordination of four participants, and management of several devices can pose considerable difficulties. The following table illustrates the obstacles we faced during this experiment and describes why some data were lost in the process.

Table 3.3 Summary of the challenges and lost data experienced for each quartet. The last column indicates whether the quartet was included in future analyses despite encountering these challenges and data loss.

Quartet*	Challenges	Lost data	Inclusion in analysis
1	The medium background noise was played at 62 dBA, rather than 54 dBA.	The between-conversations survey lacked one question (question number 2), and the video recording was briefly lost in mid-experiment.	Yes
2		Significant data loss: head motion, speech, and video data were absent for two of the six conversations	No
3&4	One participant's microphone malfunctioned, failing to record any conversation.	Although data from the high background noise condition was irretrievable, microphone data was recovered by using the data captured by the table microphone.	Yes
5	Member of HSSS staff, fitting the specified age, was appointed as a substitute for last-minute participant cancellation.		Yes
6	A participant arrived without their hearing aids but was provided with a replacement pair. These were configured by the HSSS audiologist based on the participant's latest audiogram.		Yes
12	The experiment initiated the aided condition with one IH participant not wearing their hearing aids, requiring an extra conversation in the medium background noise and the inclusion of another topic (Appendix C).		Yes
13	One participant declined to interact with the slider during the experiment but expressed a willingness to complete the study		Yes
14	Member of HSSS staff, fitting the specified age, was appointed as a substitute for last-minute participant cancellation.		Yes

* Quartets 7, 8, 9, 10, 11, 15, 16, 17, 18, and 19, although not represented in the table, all provided complete data and were included in subsequent analyses.

Last-minute cancellations were addressed by substituting absent participants with members of the HSSS. This happened twice (Table 3.3), potentially influencing these particular quartets in distinct ways. In Quartet 5, an HSSS member replaced a participant with impaired hearing, simulating hearing loss by using earplugs in conditions without hearing aids. While this approach may have worked for him, the other participant with impaired hearing felt isolated, as they were the only one experiencing genuine hearing loss. In Quartet 14, the HSSS member replaced a participant with normal hearing, closely matching their own hearing ability and age, leading to no noticeable differences from the perspective of the other participants. To ensure these substitutions did not impact the overall findings, all analyses in this thesis were conducted both with and without these two quartets.

3.5.3 Final sample description

The data analysis involves a total of 72 participants (39 females), who are categorized into two groups: those with normal hearing (n=36) and those with impaired hearing (n=36). All participants were fluent in English, aged between 50-78 (mean age=67.86 years), and were residents of Glasgow, UK.

3.5.3.1 Hearing Abilities Assessment

The normal hearing group had a mean better ear average (BEA) of 13.15 dB HL, ranging from -2.5 dB to 28.75 dB HL. In contrast, the impaired hearing group exhibited a mean BEA of 50.15 dB HL, ranging from 35 dB to 65 dB HL. These scores denote a clear distinction between the two groups in terms of pure tone thresholds aligning with World Health Organisation criteria for normal hearing and moderate hearing loss. Moreover, the division between the groups is further highlighted by the discrepancies in their perceived hearing handicap, as well as spatial and qualitative aspects of hearing. The normal hearing group had an HHIA/E mean of 14.4, and an SSQ12 mean of 6.8, while the impaired hearing group had an HHIA/E mean of 42.9 and an SSQ12 mean of 4.6.

3.5.3.2 Descriptive Questionnaires

The results show two distinct groups in terms of hearing ability status, but uniformity across groups in terms of overall level of positive affect, autistic traits, empathy, and loneliness.

The Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988) offers a range from 10 (low positive/negative affect) to 50 (high positive/negative affect). Results revealed that participants had a higher mean Positive Affect (34.38) compared to Negative Affect (18.61), with no significant differences ($p = >.05$) found between PwNH and PwIH. This outcome indicates that our sample is more likely to be represented by high energy, concentration, and pleasurable engagement, and less by aversive mood states, such as anger, contempt, disgust, and guilt.

Regarding the Autism Spectrum Quotient Short version (AQ-10 adult) (Baron-Cohen et al., 2001; Hoekstra et al., 2011) the maximum possible score is 10, with a widely accepted threshold of 6 which is considered suggestive of autism spectrum disorder traits. Participants obtained an overall mean of 2.66, with no differences ($p = >.05$) between PwNH and PwIH. 66/72 participants scored below the threshold of 6 showing a low chance of autistic traits. The remaining six were equally distributed between the two groups (PwNH = 3 and PwIH = 3), and each was part of a different quartet (3, 4, 6, 9, 15, 19).

Empathy was measured with the Interpersonal Reactivity Index (IRI) (Davis, 1980) which has a score for each subscale that can vary from 0 to 28. While there are no established thresholds for what would constitute a "low" or "high" level of empathy on this scale, higher scores on each of the subscales indicate higher levels of that specific dimension of empathy. Overall, the participants yielded a total mean score of 53.12, with subscale means of 22.17 for the Empathy subscale, 18.33 for the Perspective Taking subscale, and 12.57 for the Fantasy subscale. There were no differences ($p = >.05$) in these scores between PwNH and PwIH.

Lastly, the UCLA Loneliness Scale (Russell, Peplau & Ferguson, 1978) was applied. This tool, which has scores ranging from 0 to 60, serves as a measure of loneliness, with higher scores indicating a greater degree of

perceived loneliness. In this study, participants averaged a score of 11.61, suggesting that feelings of social isolation were generally not a concern within this sample. No differences ($p = >.05$) were found between PwNH and PwIH.

3.6 Conclusion

Multimodal synchronous data was collected from 72 PwNH and PwIH during face-to-face group conversations in different levels of background noise.

The chapter outlines a complex laboratory study that sets the stage for the analyses presented in the upcoming chapters. These analyses include 18 quartets involved in this study, including both PwNH and PwIH. The study's primary objective was to explore participants' self-perception of conversation success across three diverse background noise conditions and to probe the influence of hearing aids within these scenarios. The data collection procedure yielded 108 video-recorded conversations, audio, motion tracking data, and continuous participant feedback. This chapter describes the participant recruitment, the experimental design and the procedure detailing every step of collecting data from session one to the debriefing sessions. The results show that participants engaged in an authentic experience, paralleling the levels of difficulty they encounter in real-life scenarios outside the lab. However, their feedback indicated certain limitations of the current study, such as concern over the realism of the conversation topics. Additional limitations, as described in the challenges section, included both human error and technical hiccups leading to some data loss. In response to some of the challenges we encountered, we found ways to rectify issues during the experimental day, coming up with in-the-moment solutions.

The resultant dataset paves the way for further analysis aimed at understanding conversation success across varied hearing conditions, bringing us closer to finding ways to operationalise the concept of conversation success.

Chapter 4
Self-reported perceptions of conversation success (Study 2)

Parts of this chapter have been presented in the posters/proceedings papers mentioned below. While some of the figures are reused, most of the content presented here is original.

- Poster, ISAAR workshop, 2023: Nicoras, R., Fischer, R., Hadley, L.V., Smeds, K., Godfrey, M., Buck, B., & Naylor, G. (2023) Conversation success in small groups, ISAAR satellite workshop on communication, August, Copenhagen, Denmark
- Forum Acusticum conference proceedings: Nicoras R., Hadley LV., Smeds K., Fischer R., Godfrey M., Buck B., Naylor G. (2023), 'Small group conversations: communication behaviour and success', Proceedings of the Forum Acusticum, Torino, Italy, Sep 12th, 2023, A10-02

One of the analyses (Ordinal pattern analysis) in this chapter has been conducted in collaboration with Dr. Timothy Beechey who developed the R package for this analysis.

4.1 Chapter Summary

This chapter explores in-the-moment perception of conversation success viewed by PwNH and PwIH while they are engaging in group conversations. Based on the method described in Chapter 3, here I demonstrate that background noise, hearing aids and hearing status can influence how individuals perceive conversation success when involved in group conversations.

Firstly, I focus on ways to measure conversation success and discuss the potential strengths and limitations of using self-reported ratings. Then, I focus on 3 research questions (RQ) and their corresponding analyses based on

participants' self-reported ratings. RQ1 is concerned with how individuals rate conversation success during versus after the conversation. RQ2 explores how the seven themes of conversation success (Chapter 2) relate to the overall perception of conversation success. RQ3 investigates how background noise, hearing aids and hearing status impact participants' ratings of conversation success. In this chapter, I also address the limitations of ordinal scales from a statistical perspective. Although widely analysed in traditional frequentist statistics, often through non-parametric testing (group-level), I also used less traditional statistical alternatives that might better fit this type of data (individual-level).

4.2 Background

Various measures have been proposed to assess potential indicators of success in conversational studies, including the efficiency of information exchange observed in tasks like storytelling (Ramsberger & Rende, 2002) or Diapix/map/maze tasks (Foltz et al., 2015; Garrod & Pickering, 2004), speaker-listener alignment (Garrod & Pickering, 2004; Holler & Wilkin, 2011; Stephens et al., 2010) the ability to anticipate the end of a speaker's turn (De Ruiter et al., 2006; Hadley et al., 2020) and conversation fluency (Lind et al., 2006; Tye-Murray, 2003). While these studies provide insights on how to measure different aspects of a conversation, they predominantly rely on objective indicators and overlook the subjective perceptions of interlocutors. That is why interlocutors' self-reported perception of success is the focus of this chapter.

In psychology, perception usually refers to the human ability to interpret, analyse, and give meaning to the sensory information they receive from the environment. Nonetheless, when it comes to abstract concepts such as good and bad or success and failure, the term "perception" takes a broader meaning. In this context, perception refers to the cognitive process that makes sense of abstract information and ideas based on personal experiences, beliefs, cultural background, and cognitive biases. But while asking participants removes the issue of inferring from their behaviour, reporting perception refers to the expression of an internal evaluation and brings another set of potential

biases. For instance, social desirability bias implies that individuals might respond in a way they think would make them socially accepted. Other potential biases would be the response bias where the way questions are framed can influence how perceptions are reported or memory bias when memories of perceptions can change and what people report might not accurately reflect their original perception. While no experiment assessing self-perception can fully account for all potential biases, self-reported perception provides valuable information despite its subjectivity. Asking people what they think about a topic and how they perceive an event opens the door to a realm of personal experiences. Self-reported measurements are usually assessed retrospectively, after the event. However, face-to-face conversation is an "in-the-moment" action. It unfolds "here and now" and any assessment outside of this time and space can result in lost information. Moreover, conversation is a continuous process, it unrolls in time, therefore the self-perception related to a conversational variable can have different values across that time. This is why the current study aims to measure participants' perception of conversation success both during and after the conversation.

Contrasting in-the-moment and retrospective self-reported ratings is a popular practice, especially in the EMA literature. In the emotion literature, the accessibility model of emotional self-report (Robinson & Clore, 2002) also supports the distinction between in-the-moment and retrospective self-reported ratings, arguing that momentary reports reflect feelings based on present experiential knowledge and cannot be relived once the emotional episode ends. In contrast, retrospective assessments involve recalling emotions associated with specific events or over a defined time period. For example, several EMA studies show that in-situ reports can be complemented by retrospective assessments rather than replaced (Leertouwer et al., 2022; Lelic et al., 2023; Wu et al., 2020). However, in our study, the period of time between the in-the-moment assessment (completed using a slider during the conversation) and the retrospective assessment (completed through a survey after the conversation) is very short, which might lead to similar ratings of conversation success. RQ1 explores the contrast between in-the-moment and retrospective assessments by

correlating ratings of conversation success during the conversation with ratings of success after the conversation.

In Study 1 (Chapter 2), seven unique themes related to conversation success were identified: (1) *Being able to listen easily;* (2) *Being spoken to in a helpful way*; (3) *Being engaged and accepted*; (4) *Sharing information as desired*; (5) *Perceiving flowing and balanced interaction*, (6) *Feeling positive emotions*; (7) *Not having to engage coping mechanisms*. However, there is limited understanding of how these themes interrelate and how they could be perceived during face-to-face in person conversation. While the first two themes were rated as more important than the others, it remains unclear which of the seven themes are most strongly associated with the in-the-moment perception of conversation success. Therefore, in this study, interlocutors received a survey integrating questions reflecting each theme (see Table 4.1) after each conversation. RQ2 explores how the survey questions interrelate to gain insights into the underlying relationships among these conversation success themes.

Chapter 3 provided a detailed description of how the current study aimed to manipulate conversation success introducing adversities such as variations in background noise levels (30 dBA, 54 dBA, and 72 dBA) and requiring participants with impaired hearing to remove their hearing aids in half of the conversations. The conversations in low noise level at 30 dBA, which is likely inaudible for most listeners, were expected to be perceived as more successful. The conversations in medium noise level at 54 dBA, resembling a moderately populated social environment, were anticipated to be perceived as less successful compared to the low noise level. The conversations in high noise level at 72 dBA, equivalent to noise levels found in a medium-occupancy restaurant (To & Chung, 2015), were expected to be perceived as the least successful. Given that communication is more challenging for PwIH (Kiessling et al., 2003) and they tend to report difficulties in group settings (Vas et al., 2017) we anticipated that PwIH would perceive lower conversation success than PwNH. However, studies suggest that when provided with hearing aids, PwIH behave more similarly to PwNH (E. B. Petersen et al., 2022);

therefore, we anticipated that they would perceive higher levels of conversation success when aided.

This chapter addresses how individuals rate conversation success during versus after the conversation (RQ1), how their perceived success relates to the seven themes of conversation success found in Chapter 2 (RQ2) and how adversity (background noise, hearing aid usage and hearing status) influenced their ratings of conversation success. The latter uses the self-reported ratings collected during and after the conversation as highlighted in Figure 4.1.

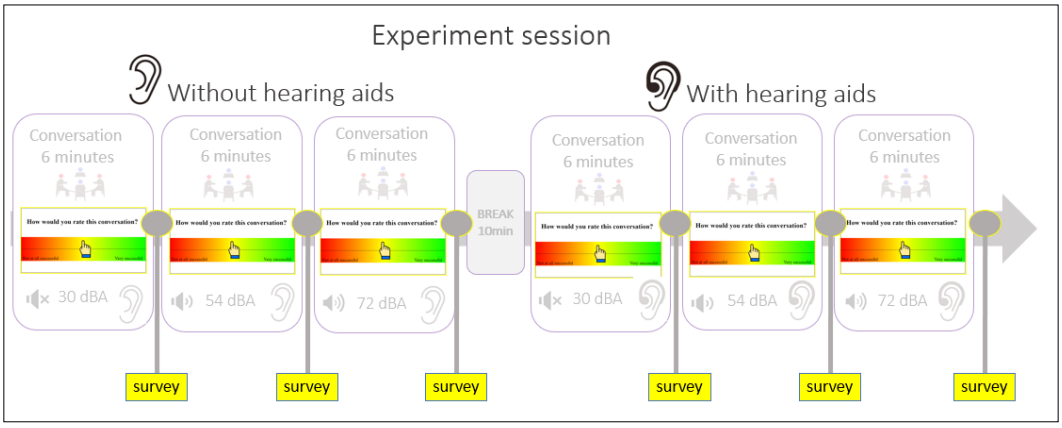


Figure 4.1 Description of the experimental session: six conversations corresponding to six conditions.

This study integrates three distinct methods for gathering self-reported perceptions of conversation success: 1) continuous recording of participants' perceptions of success throughout the conversations using a slider; 2) evaluation of overall perceived success after each condition (survey item 1, Table 4.1); 3) assessment of specific conversation success themes after each condition. To be more specific, we translated the seven facets of conversation success (Chapter 2) into a survey (Table 4.1). First, we had the question related to the overall perceived success. Then, on the premise that difficulty in communication can trigger communication effort (Beechey et al., 2020), we introduced an additional item (2) that was not reflected in the facets of conversation success. Communication effort refers to all the actions participants believe they must take in order to ensure the success of a conversation, actions that may not typically be required in the absence of the

adversity. The following items 3-9 each focused on a different facet of conversation success (Chapter 2). Item 4 which reflects the theme "Being able to listen easily", (Chapter 2) was split into three (4a, 4b and 4c) corresponding to each interlocutor, but will be analysed together as one mean (Item 4abc) in the analysis.

Table 4.1 The survey administered after each condition.

Motivation	Survey item	Rating scale
Overall conversation success	1) To what extent would you rate this conversation as successful?	1 = unsuccessful 5 = successful
Communication effort (Beechey, 2020)	2) Did you do anything or say anything in a particular way to make this conversation successful?	1 = not at all 5 = very much
Seven facets that drive conversation success (Chapter 2)		
Being able to listen easily	3) How easy was it for you to follow the conversation?	1 = not easy 5 = very easy
Being spoken to in a helpful way	4a) To what extent was the person sitting on your left speaking in a helpful way? 4b) To what extent was the person sitting in front of you speaking in a helpful way? 4c) To what extent was the person sitting on your right speaking in a helpful way?	1 = not helpful 5 = very helpful
Being engaged and accepted	5) How connected did you feel with other participants?	1 = not connected 5 = very connected
Sharing information as desired	6) To what extent did you share information successfully?	1 = not successfully 5 = very successfully
Feeling positive emotions	7) How enjoyable was this conversation for you?	1 = not enjoyable 5 = very enjoyable
Perceiving flowing and balanced interaction	8 To what extent was this conversation flowing smoothly?	1 = not smoothly 5 = very smoothly
Not having to engage coping mechanisms	9) How often did you feel uncomfortable/anxious during this conversation?	1 = never 5 = always

4.3 Aims

The main aim of this study is to investigate how adversity affects the perception of conversation success in face-to-face group conversations.

However, we also engage in exploratory investigations to broaden our understanding (RQ1 and RQ2). First, we compare the perception of conversation success in different moments: during the conversation versus after the conversation (RQ1). Then we explore the ratings given to the survey items to investigate how the seven themes of conversation success relate to participants' perceived success in face-to-face interaction (RQ2). Finally, we investigate the impact of background noise, hearing aid usage and hearing ability on interlocutors' ratings of conversation success (RQ3).

Research questions:
RQ1. How do in-the-moment and retrospective conversation success ratings relate to each other?
RQ2. Do multiple factors contribute to the overall perception of success?
RQ3. How do noise level, aiding and hearing ability affect perceived conversation success?

I address each RQ in turn due to the distinct analyses I uses to address them.

4.4 RQ1. How do the in-the-moment and retrospective conversation success ratings relate to each other?

Hypothesis: Ratings of conversation success during and after the conversation are mutually consistent for all participants.

4.4.1 Analysis

Data used to test this hypothesis were the measurements throughout the conversation, and the post conversation overall success question: "To what extent would you rate this conversation as successful?". Data from the slider consists of a timeseries of data reflecting success as rated via the slider, throughout the conversations. Notably, participants interacted with the slider in different ways. For instance, some people slid their finger on the screen creating a series of datapoints between the previous position and the current one, while other people used the tapping method creating only two datapoints

for the same change in rating. Similarly, some people used the entire range of the slider, while others used small ranges. Finally, some people simply forgot to interact with the screen. The slider software (Emotouch (Louven et al., 2022)) recorded every change made to the slider by participants. When participants were not actively touching the slider with their finger, no data was recorded during those periods. Therefore, we filled in the missing values ('fillmissing' function in MATLAB) for each timestamp such that we can see the position of the slider during the entire six minutes of conversation. (Figure 4.2) for all participants regardless their method of interaction with the slider (i.e.: tapping or sliding).

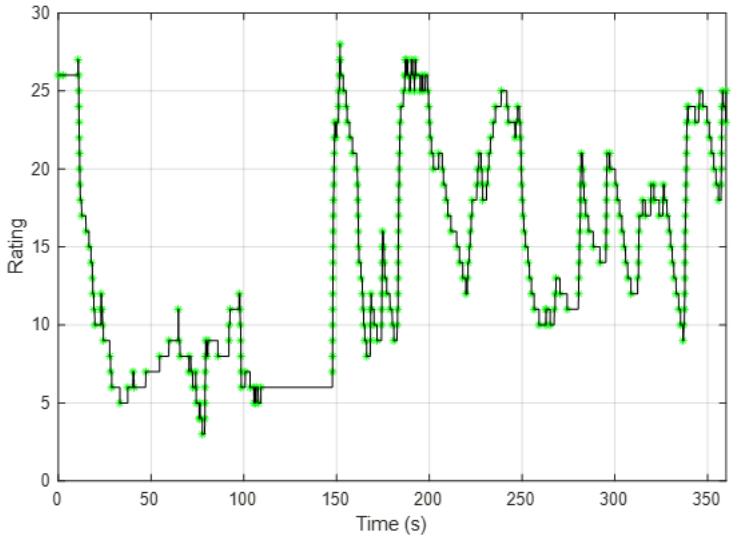


Figure 4.2 Data from one participant in one conversation showing the ratings from the slider during the conversation (green stars) as well as the ratings obtained by using the "fillmissing" function (black line).

Pearson correlation was employed to compare scores from the in-the-moment slider and post-conversation feedback on overall success (survey question 1). We examined correlations derived from several metrics: mean of data post-processing (after gap-filling), mean scores from different intervals (every minute), and separate means from both the highest and lowest scores recorded**.** Bonferroni correction was applied.

4.4.2 Results

The results showed a strong correlation between the mean collected through the in-the-moment slider and the ratings of conversation success collected through the post-conversation survey question 1. Also, the analysis explored different periods during the conversation (every minute) as well as periods with the highest and lowest mean ratings of conversation success (Table 4.2):

Table 4.2 Pearson correlations between the data collected through the slider during the conversation versus item 1 from the survey administered after the conversation.

Slider means versus Item 1	**All (filled)**	**Min 1**	**Min 2**	**Min 3**	**Min 4**	**Min 5**	**Min 6**	**MAX**	**MIN**
Survey item 1	.73**	.54**	.63**	.66**	.70**	.73**	.77**	.67**	.71**

** Correlation is significant p <.01
* Correlation is significant p < .05

4.4.3 Discussion

A novel aspect of this study was the assessment of success conducted both during, and after the conversation. While the data collected during the conversation points out moments in time when participants perceived a change, survey item 1 assesses participants' perspective on the conversation overall. The results showed that participants were consistent in their ratings of conversation success regardless of the time of the assessment. However, the highest correlation coefficient was found between the slider data collected in the last minute of the conversation and the ratings after. This might be because the two assessments are closer in time. The strong correlation between survey question 1 and the ratings during the conversation provides support for considering the inclusion of the survey question 1 in the subsequent analyses where we investigated the effects of the background noise, hearing aids and hearing group on the perception of conversation success.

Both ratings (during and after) are done in the presence of the other members of the group. Participants could talk and see each other when completing their self-perception ratings. Although they could not see the details on each other's screen, noticing that another person interacted with the screen might have influenced the individual behaviour (e.g.: Person B noticed that Person A moved the slider. Consequently, Person B interacted with the slider as well.).

4.5 RQ2. Do multiple factors contribute to the overall perception of success?

Hypothesis: The overall perception of success includes seven distinct factors relating to those found in chapter 2.

4.5.1 Analysis

Data to test this hypothesis are collected through the post-conversation survey. The survey consisted of a set of items designed to measure different aspects of conversation success. Participants were asked to rate each item based on their experiences during the conversation. To identify the underlying structure of conversation success, Spearman correlations were employed to examine the relationship between the survey items. Then, exploratory factor analysis (EFA) was used to investigate in more detail how the individual survey questions were structured. EFA is a statistical technique that aims to uncover latent factors that explain the patterns of correlations among observed items. It allows us to reduce the data into a smaller set of interpretable factors, facilitating a deeper understanding of the phenomenon. In this study, principal factor analysis, a common approach in EFA, was used to extract the factors. The eigenvalues of the factors were examined to determine the number of factors to retain. The Kaiser criterion (Eigenvalues > 1.0) was used to retain factors that accounted for a substantial amount of variance in the data. Upon extracting the factors, factor loadings were examined to understand the relationships between the items and the identified factors. The interpretation of the factor was based on the items with high loadings, indicating that they are strongly associated with that factor.

4.5.2 Results

In order to test if the themes of conversation success manifest in face-to-face group conversations, the correlation between the survey item 1 asking about the overall success (How successful would you rate this conversation?) and the rest of the items relating to individual themes was applied using Pearson test. We are interested in how each item of the survey correlates with the item number 1 (Table 4.3).

Table 4.3 Pearson correlation coefficients for all survey items against survey item 1 and each other:

Item 1	Item 1 Overall success	Item 2 Effort	Item 3 Easy listening	Item 4 Helpful speaker	Item 5 Engaged and accepted	Item 6 Sharing information	Item 7 Flowing and balanced	Item 8 Positive emotions
Item 2 Effort	**-.05**							
Item 3 Easy listening	**.33****	.00						
Item 4 Helpful speaker	**.30****	-.03	.36**					
Item 5 Engaged and accepted	**.26****	-.05	.16**	.22**				
Item 6 Sharing information	**.11***	.05	.07	.09	.00			
Item 7 Flowing and balanced	**.33***	.02	.43**	.47**	.18**	.17**		
Item 8 Positive emotions	**.18****	.00	.13**	.17**	.06	.19**	.25**	
Item 9 No coping mechanisms	**.10***	-.05	.24**	.15**	.08	.01	.15**	.16**

** Correlation is significant p <.01
* Correlation is significant p < .05

Among all the items, item 2 (related to effort) stands out as the only one that does not show a significant correlation with the question regarding overall success (Item 1). This observation confirms that item 2 does not align with the seven themes of conversation success (Chapter 2). While items 3-9 are based on the facets shown to drive conversation success (Chapter 2), item 2 is meant to cover the effortful communication aspect.

Next, the exploratory factor analysis was conducted only on the items that showed a significant correlation with the overall success item 1 i.e., the seven themes of success, to address their interrelationships. Given the non-normal distribution of the data, the Principal Axis Factoring method was chosen for extraction (Osborne et al., 2011). The analysis resulted in one factor with Eigenvalues > 1, explaining 48% of the variance collectively. The result

was confirmed by both Parallel Analysis (Vivek et al., 2007) and the scree plot (Figure 4.3), suggesting the presence of only one underlying component.

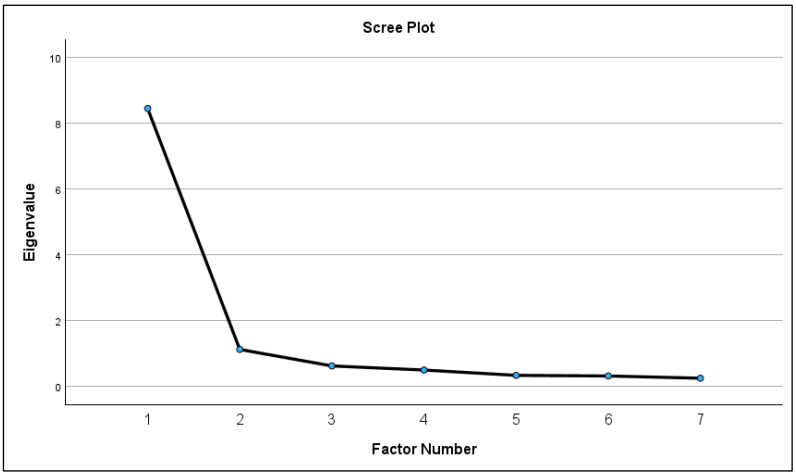


Figure 4.3 Scree plot showing the elbow point appearing after one factor.

Table 4.4 Factor loadings for each survey item. Factor loadings greater than 0.3 indicate a strong relationship between the item and the corresponding factor.

Items	Factor loadings
	1
3) How easy was it for you to follow the conversation?	.833
4) To what extent have you been spoken to in a helpful way?	.809
5) How connected did you feel with other participants?	.873
6) To what extent did you share information successfully?	.884
7) How enjoyable was this conversation for you?	.918
8) To what extent was this conversation flowing smoothly?	.855
9) How often did you feel uncomfortable/anxious during this conversation?	.575

4.5.3 Discussion

We previously found that PwNH and PwIH identified seven themes of conversation success (Chapter 2). Here I demonstrated that when these were translated into individual questions (item 3, 4, 5, 6, 7, 8, and 9) on a 5-point Likert-scale and administered after actual conversations, they form only one factor. Factors were extracted using the principal axis factor method, though note that other extraction methods available in SPSS (PCA, unweighted least squares, generalized least squares, maximum likelihood, alpha factoring, and

image factoring) showed the same outcome. The one factor found could be interpreted as conversation success, but it could also mean something else. For instance, the item 7 ('How enjoyable was this conversation for you?') appears to be more significantly correlated with other items than any other item and has the highest loadings on the resulting factor. This suggests that the one factor identified as relating to overall conversation success could be potentially more related to enjoyment than to success itself. Items 6, 8 and 9 were least correlated to the overall perception of conversation success. Items 6 (*Sharing information successfully*) and Item 8 (*Perceiving flowing and balanced interaction*) relate to informational conversations. A potential explanation could be that the topics used in this experiment did not follow a clear goal, hence participants did not have a specific task to complete.

This study was not designed to validate a potential conversation success scale, but rather to explore how people perceive the facets that have been shown to drive conversation success in face-to-face group conversations. The items might be correlated because they measure similar concepts or might be correlated because they are based on repeated measurements from the same person consecutively. This study could serve as a pilot experiment for future research aiming to develop a multi-factor measurement of conversation success.

Given the strong correlations of all seven items with the overall ratings of conversation success, we include survey item 1 as the measure of conversation success in the subsequent analyses (RQ3).

4.6 RQ3. How do noise level, aiding and hearing ability affect perceived conversation success?

Hypothesis 1 (H1): Conversation success decreases as background noise increases.

Hypothesis 2 (H2): When PwIH are not using hearing aids, conversation success is lower for both PwNH and PwIH.

Hypothesis 3 (H3): In any given condition, PwIH are likely to perceive a reduced level of conversation success compared to PwNH.

The hypotheses were tested from two statistical perspectives. Firstly, I described the group-level analysis, which employs a traditional approach of testing against the null hypothesis and presents results based on the mean of the sample. These conventional methods are instrumental in deriving group-level inferences and forecasting for a hypothetical average individual. Although they provide insights into the dominant trends of perceived conversation success, they may overlook other data tendencies. Another constraint with such analyses is measurement precision. Unlike tangible metrics, conversation success, similar to many psychological constructs, lacks a definitive scale. It is challenging to assert, for instance, that the perceptual gap between ratings of 1 and 2 is equivalent to that between 4 and 5. To better capture these individual differences, our assessment also incorporates individual-level analyses designed for ordinal data.

4.6.1 Group-level analysis

4.6.1.1 Methods

The hypotheses were assessed using nonparametric tests, focusing on the overall post-conversation perception of success (item 1, Table 4.1). Friedman tests were used to examine the primary effects of background noise levels on the perception of conversation success in both aided and unaided conditions. To determine differences between groups, post-hoc Whitney - Wilcoxon tests were used (SPSS software).

4.6.1.2 Group-level results

Friedman tests were conducted to determine if the perception of conversation success was affected by background noise levels (H1) across both aided and unaided conditions across both groups (PwNH and PwIH). Results (figure 4.4) showed significant differences between the three levels of background noise $\chi2(df = 2, N = 71) = 60.41$, $p < .05$, and $\chi2(df = 2, N = 71) = 40.62$, $p < .05$. Follow-up pairwise comparisons using Whitney - Wilcoxon signed rank tests showed a significant drop in conversation success ratings in high background noise levels, compared to both medium and low background noise levels ($p< .001$). In the aided conditions, PwNH and PwIH perceive conversation success similarly ($p = >.05$). In the unaided conditions, PwIH

perceive lower conversation success than PwNH in 30 dBA (p = < .05) and 54 dBA (p = < .05) (H2 and H3).

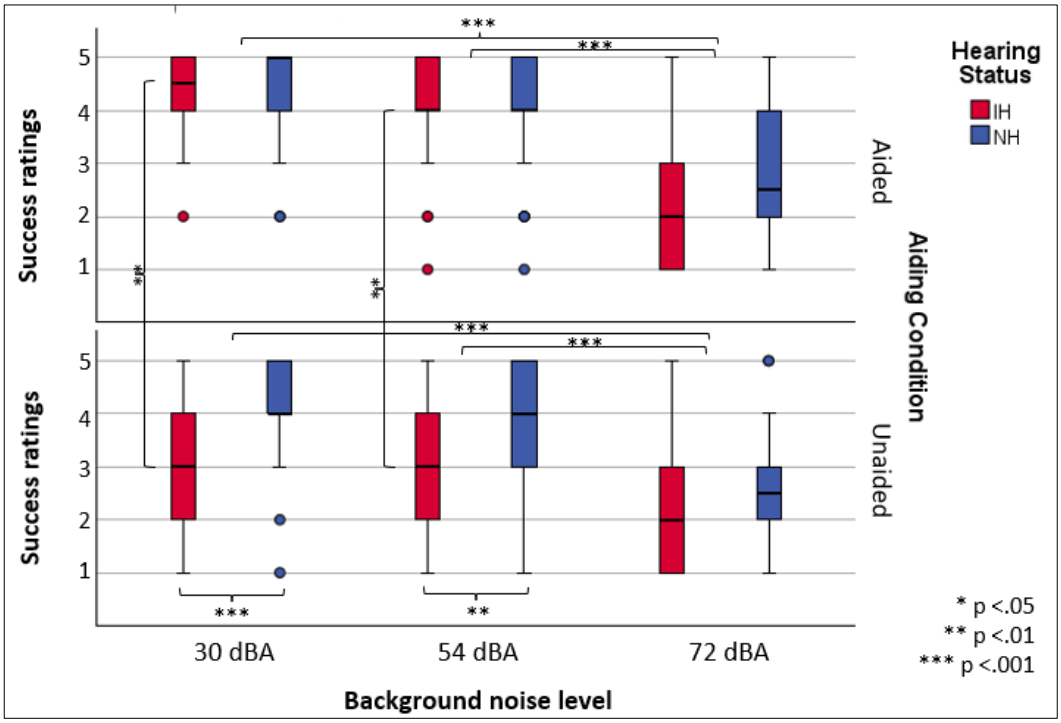


Figure 4.4 Boxplots showing the ratings of conversation success (y axis) against the background noise level (x axis), considering the hearing status (blue = PwNH, red = PwIH) and the hearing aids (upper side = with hearing aids, lower side = without hearing aids).

4.6.2 Individual-level analysis

4.6.2.1 Methods

To further explore the above group-level analysis, I conducted a series of Observation Oriented Modelling analyses. These are a statistical modelling approaches that consider events from the perspective of individual observations (Beechey, 2023). This approach primarily revolves around building models based on individual observations or cases and then assessing how well these models fit with the observed data.

H1 was tested using Ordinal Pattern Analysis using the OPA Package in R (Beechey, 2023). This is a method designed to test the extent to which the

observed data match a hypothesis of relative change. The relative change is defined by the relationship between the observed data in different conditions. These can take three forms: increase (↑), decrease (↓) and equality (=). As the method and interpretation of OPA may be unfamiliar, here follows a description based on a small example (adapted from Beechey, 2023).

Example OPA

Our initial hypothesis, H1, states that conversation success ratings decrease in medium background noise (54 dBA) condition compared to low background noise level (30 dBA) and decrease in high background noise (72 dBA) condition compared to medium background noise level (Figure 4.5a). Data observed from a conversation can be seen in Figure 4.5b.

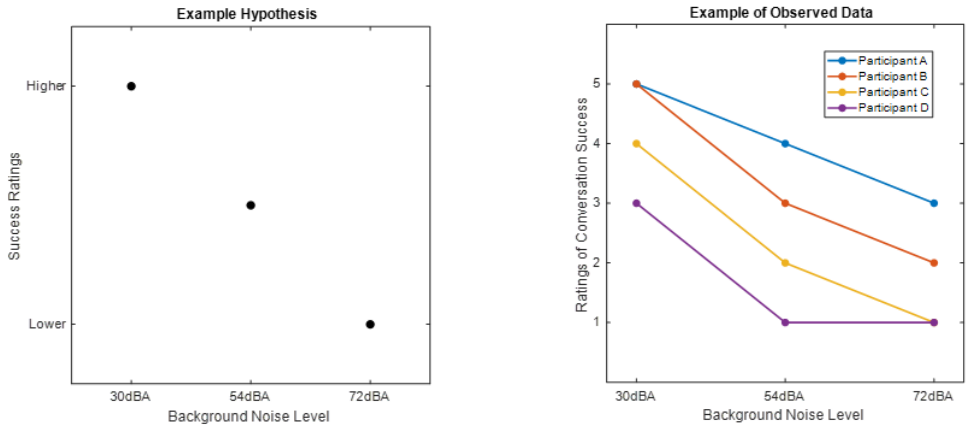


Figure 4.5 (a) Example hypothesis: 30 dBA ↓ 54 dBA ↓ 72 dBA. and (b) example of observed data.

In this case, we have the same individuals rating conversation success repeatedly, in three different conditions. Hence, there are 3 orders we can determine for each individual: (1) ratings of conversation success in 30 dBA relative to 54 dBA, (2) ratings of conversation success in 54 dBA relative to 72 dBA, and (3) ratings of conversation success in 30 dBA relative to 72 dBA. For each order that we test against our hypothesis we obtain a correct/incorrect classification. The percentage of correct classifications (PCC) is the most important statistic in OPA. The PCC of the hypothesis for each individual can be calculated simply as the number of ordinal classifications (↓, ↑, or =) which

match the corresponding ordinal classification in the hypothesis, divided by the total number of classifications, expressed as a percentage (Beechey, 2022). For the data shown in Figure 4.5b, we will obtain the following orders:

	30 dBA – 54 dBA	54 dBA - 72 dBA	30 dBA – 72 dBA	PCC
Hypothesis	↓	↓	↓	100%
Person A	↓	↓	↓	100%
Person B	↓	↓	↓	100%
Person C	↓	↓	↓	100%
Person D	↓	=	↓	66.6%

While the observed data of Persons A, B and C in the example above fit the hypothesis 100%, for Person D, two of the ordinal relations are decreases (↓), and one is an equality (=). The hypothesis consists of three decreases (↓). So, for this individual, 2/3 = 66.6% of the ordinal relations are correctly classified by the hypothesis. The second most important statistic value in OPA is the chance value (c-value). The c-value is showing how likely is to get the same PCC from other permutations. In other words, c-value is the proportion of permutations which produce a PCC at least as high as that derived from the observed data (Beechey, 2022). A c-value of 0 would indicate that the results (PCC) are attributed to the experimental conditions, while a chance value of 1 indicates that a similar result (PCC) could be obtained by many other orderings hence the results are not attributed to the experimental conditions.

4.6.2.2 Individual-level results

Ordinal Pattern Analysis (OPA) was conducted on 3 observations for 18x4 individuals, 36 in each of two groups (PwNH and PwIH). The expected pattern (Figure 4.6) is informed by the above group-level results showing no change in ratings of conversation success between 30 dBA and 54 dBA background noise, but a decrease between 54 dBA and 72 dBA background noise levels (=↓↓)

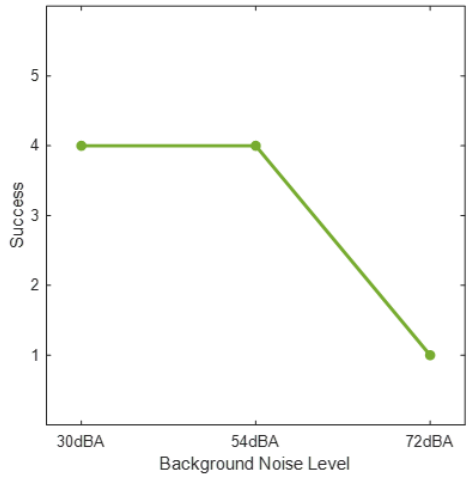


Figure 4.6 Expected pattern showing a decrease in ratings of conversation success, based on group level results (H1).

Two models were fitted, one for each condition: aided and unaided. Chance-values were calculated from 1000 random orderings.

Between subjects results for the aided condition showed an overall PCC = 65.42 (c-value = 0), showing a difference of PCC=11.86 between groups (c-value = 0.14), PwNH having a lower PCC= 59.43 (c-value = 0) than PwIH PCC = 71.30 (c-value = 0). In terms of counts 23/72 participants behaved 100% as expected.

Between subjects result for the unaided condition showed an overall PCC = 60.28 (c-value = 0), showing a difference of PCC = 3.54 between groups (c-value = 0.69), with PwNH having a PCC = 62.04 (c-value = 0) slightly higher than the PCC = 58.49 (c-value = 0) of PwIH. In terms of counts 28/72 participants behaved 100% as expected.

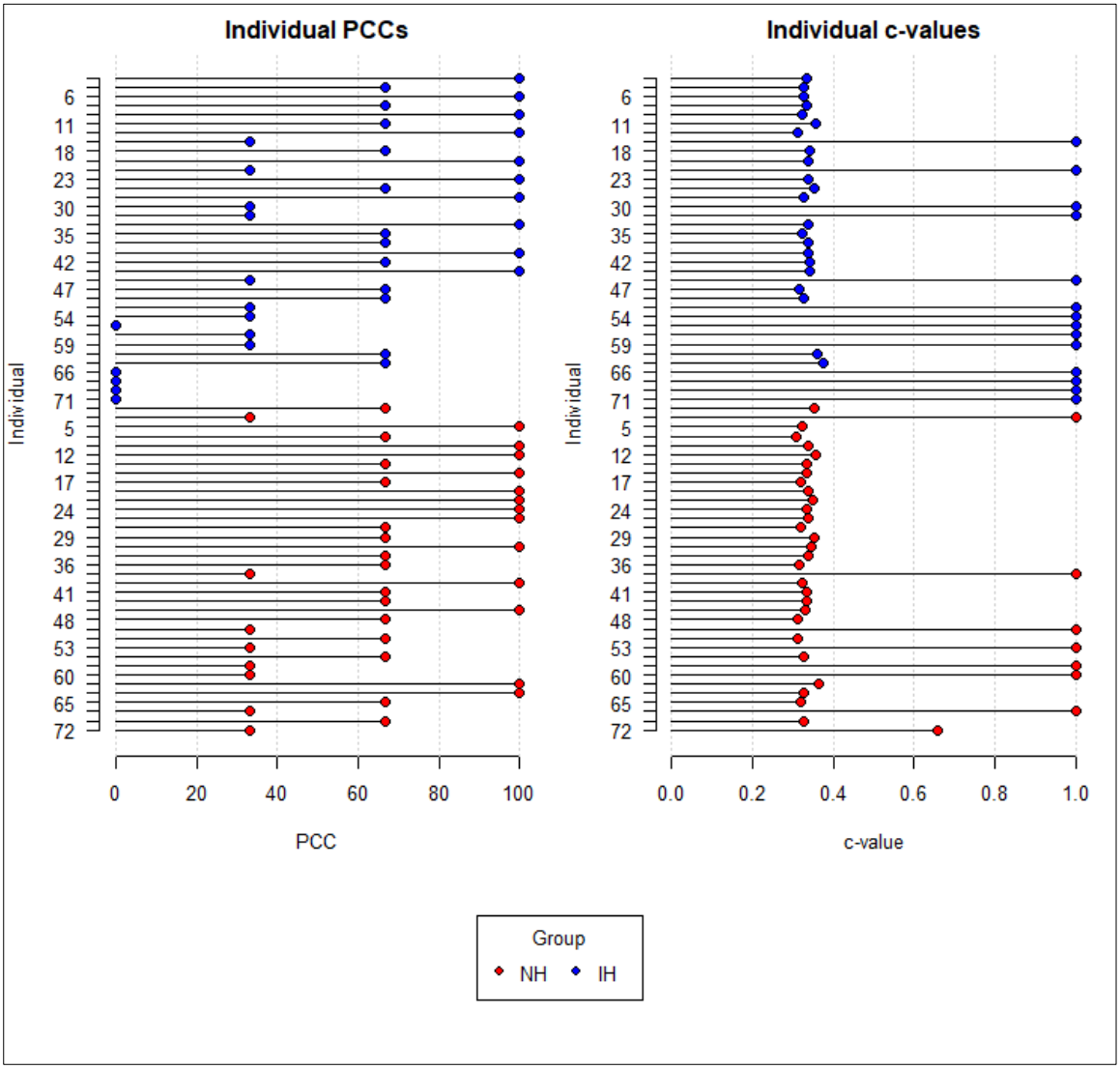


Figure 4.7 Individual PCC and C-values for the aided condition.

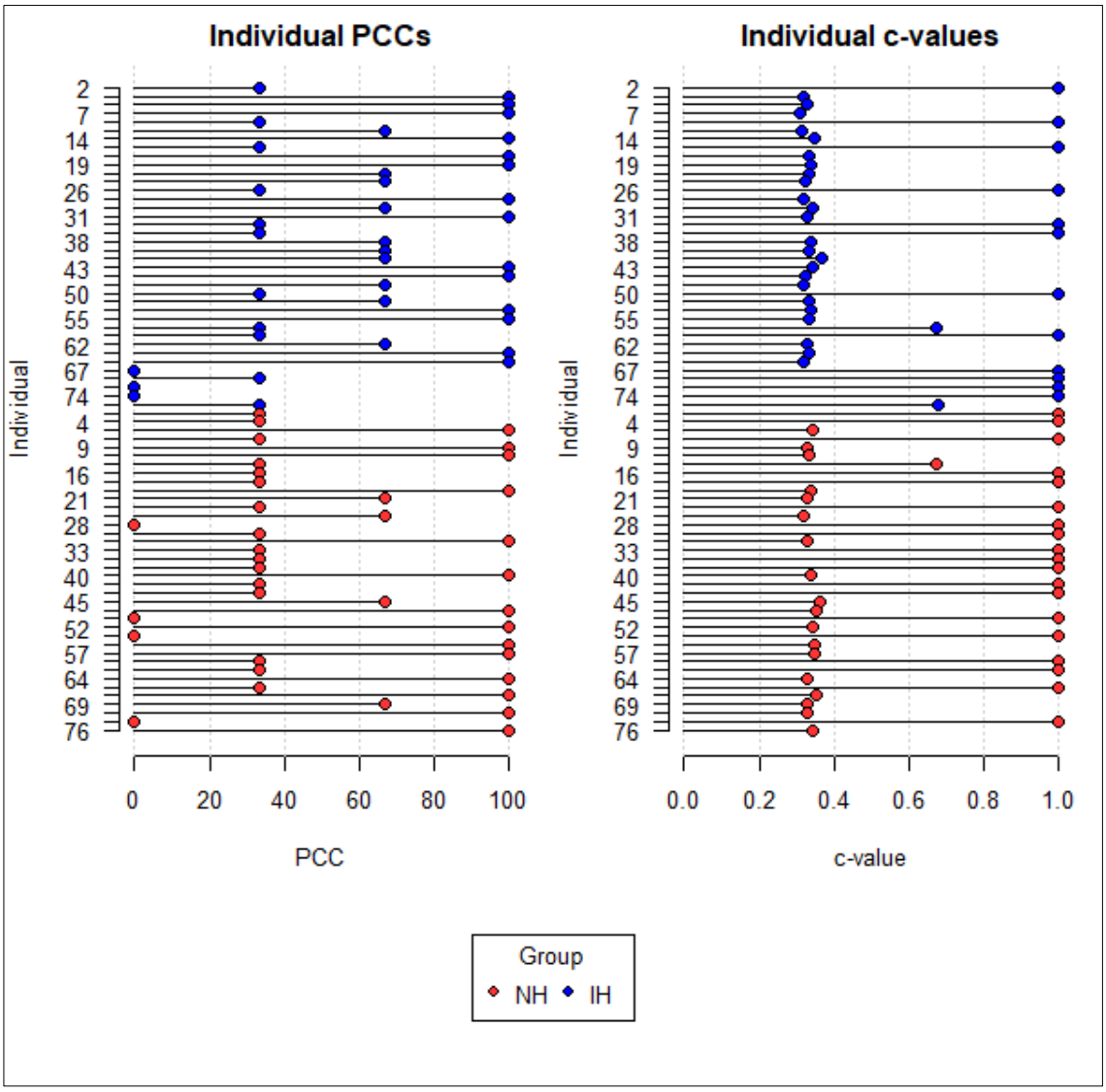


Figure 4.8 Individual PCC and C-values for the unaided condition.

4.6.3 Discussion group and individual levels

As expected, background noise level was shown to have a significant impact on the perception of conversation success for all participants. While there was no significant difference between the quiet condition (30 dBA) and the medium background noise level condition (54 dBA), participants' ratings of conversation success decreased significantly when the background noise level increased to 72 dBA. This finding supports the existing literature showing that not "*being able to listen easily*" negatively affects conversation success

(Chapter 2), and complements other work showing that high background noise relates to increased communication effort (Beechey, 2019), feelings of frustration and disturbance (Bottalico, 2022), and reduces feelings of performance success (Aliakbaryhosseinabadi, 2023). Ordinal pattern analysis (Beechey, 2022) showed that only around 60% of the observed data follows the pattern expected from the group level analysis (= ↓ ↓). However, when we investigated in-depth how many participants had ratings that matched 100% with these group-level results (= ↓ ↓), we found 23/72 in the condition with hearing aids, and 28/72 in the condition without hearing aids.

Besides the pattern proposed based on group-level results (= ↓ ↓) , another expected order could have been a monotonic decrease in ratings of conversation success as background noise increases (↓ ↓ ↓), which corresponds to our original hypothesis (H1) (i.e.: participants rated conversation success lower in 54 dBA compared to 30 dBA, lower in 72 dBA compared to 54 dBA, and lower in 72 dBA compared to 30 dBA). When we take both expected patterns into account (= ↓ ↓ and ↓ ↓ ↓), we reach a total of 78/144 participants (37/72 in the condition with hearing aids, and 41/72 in the condition without hearing aids) that behaved 100% in an expected manner. Exploratory OPA was conducted with the purpose of finding groups of participants whose ratings of conversation success do not match either of the patterns described above. While these participants do not fit our expectations, they might have behaved in a similar way between themselves. Two patterns were found: 1) pattern (= = =) suggesting no change in conversation success ratings regardless the backfround noise level and 2) patterns(↑↓↓) & (↑↓=) suggesting highest ratings in 54 dBA.

The first group (= = =) showed no change in the perception of conversation success regardless the background noise level, meaning that they rated the conversation success with the same value in 30 dBA, 54 dBA and 72 dBA. In the aided conditions, this occurred in 1 PwIH and 5 PwNH. In the unaided conditions, it was observed in 7 PwIH and once in a PwNH. This might be explained by human mistake, misunderstanding of the task, or disengagement. However, even if they did not declare to have perceived a

change, their position on the 5-point scale reflects their hearing status, PwNH tending to give higher ratings throughout the experiment, and PwIH tending to give lower ratings (Figure 4.9 - left).

The second group (patterns ↑↓↓ & ↑↓=) shows that some participants perceived higher conversation success in 54 dBA than in 30 dBA, and lower conversation success in 72 dBA than in 54 dBA (Figure 4.9-right). In the aided conditions, this occurred in 7 PwIH and 6 PwNH. In the unaided conditions, it occurred in 5 PwIH and in 4 PwNH. This unexpected behaviour might be the results of a mistake or misunderstanding of the ratings scale. However, another plausible explanation would be that people in this pattern benefited from a potential Lombard effect that appeared in the medium background noise level (54 dBA). The same effect, even if present in 72 dBA might not have provided enough compensation.

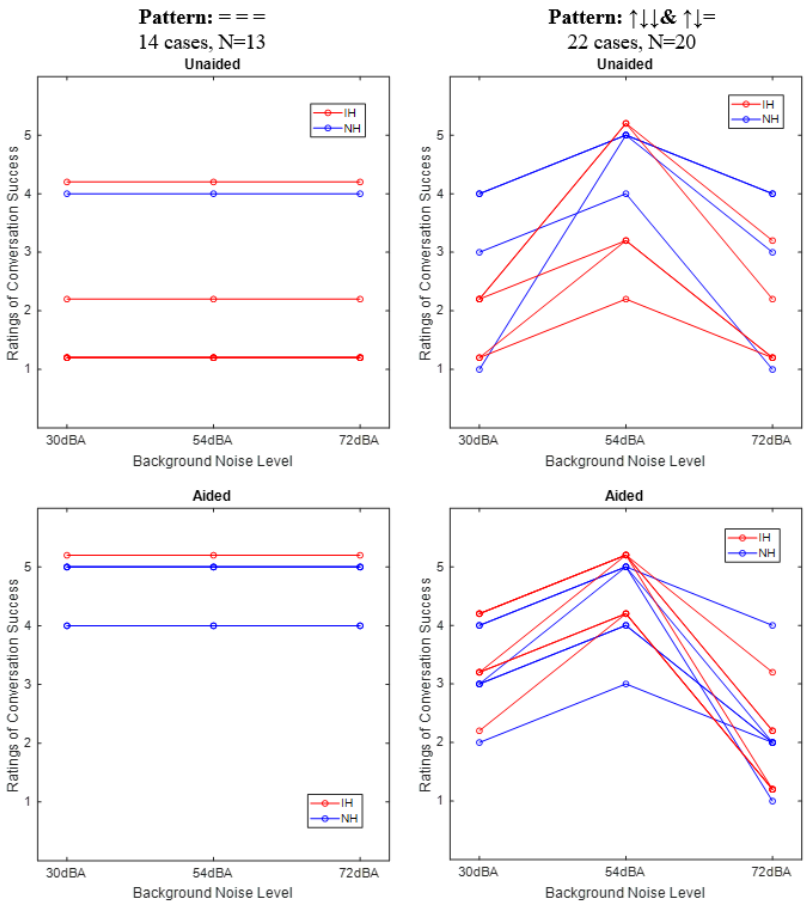


Figure 4.9 Unexpected patterns found in the observed data

The study revealed a significant improvement in conversation success for PwIH attributed to hearing aids, but in the low and medium background noise levels (30 dBA and 54 dBA). This result might suggest that participants with moderate hearing loss benefit from their hearing aids during group conversations in certain conditions. In highly adverse acoustic environments, characterized by elevated background noise levels (72 dBA), the efficacy of hearing aids in improving conversational success appears limited. However, when confronted with challenging, yet not overwhelming, acoustic conditions, participants tend to perceive a noticeable increase in conversation success when wearing hearing aids. Contrary to our expectations (H2), PwNH did not perceive higher conversation success when their interlocutors were aided.

When PwIH were unaided in low and medium background noise, their ratings of conversation success were significantly lower compared to those of PwNH (H3). However, when aided, there was no significant difference observed between NH and PwIH. This finding is consistent with existing literature suggesting that hearing aids may bridge the gap between PwIH and PwNH, making them more similar (Petersen et al., 2022). Also, when confronted with loud background noise, PwNH rate conversation success similarly to PwIH, suggesting potential challenges that both groups face in such environments.

4.7 Discussion

4.7.1 Summary of the results

In this chapter we found that the perception of conversation success during six-minute conversations is highly correlated to the perception of conversation success after the conversation ended; the themes of conversation success found in the first study (Chapter 2) are structured as one factor according to the exploratory factor analysis. Some manipulations of adversity, including alterations in background noise levels and non-usage of hearing aids, reduced the perception of conversation success. However, these outcomes did not necessarily align with our initial expectations, as detailed below.

Background noise contrast: Friedman tests revealed that background noise level had a significant effect on the perception of conversation success,

showing conversation success to decrease in 72 dBA compared to lower levels. However, no significant differences were found between 30 dBA and 54 dBA. On an individual level, only around 60% of the observed data match these group level results, leaving room for further investigation. Potential explanations are shown in an exploratory analysis presented in the section 4.6.3.

Hearing aid contrast: Hearing aids elicited a significant improvement only for PwIH in 30 dBA and in 54 dBA.

Hearing status contrast: Results showed that PwIH perceived conversation success significantly lower than PwNH, but only when PwIH are unaided in 30 dBA and 54 dBA.

4.7.2 Strengths and limitations

This study, to the best of my knowledge, is the first to assess self-perceptions of conversation success in interactions. Also, it extends beyond the usual dyadic or triadic configurations by involving four-way interactions. A significant strength lies in our novel approach to assessing conversation success both during and immediately after the conversation. Differing from previous research, our study examines ratings in real-time during the conversation. By examining these two temporal perspectives, we gain a better understanding of how participants' perceptions during the conversation relate to their post hoc evaluations. This dual assessment approach offers a holistic view of conversation success and provides insights into the potential impact of immediate impressions versus reflective evaluations. Furthermore, the analyses incorporate innovative approaches that do not assume that ordinal data can be treated as normally distributed. We conducted analyses at both group and individual levels, showcasing how these complementary analyses can yield a deeper understanding of the data. This multifaceted approach enhances the robustness of our findings and allows for a more detailed interpretation of the results.

This study also bears certain limitations that should be considered. One important challenge relates to the act of asking participants to rate conversation success while they were actively engaged in conversation. This approach may

have altered their experience by heightening awareness of the task at hand. This awareness might introduce bias into their upcoming conversations and subsequent ratings. Also, the survey presented to participants between conversations may have prompted them to be more attentive to specific aspects of their interactions, thereby affecting their subsequent responses and perceptions. Another limitation relates to the physical arrangement of participants around the table, with tablet screens displaying the sliders facing each participant. It is possible that participants could have seen the hand movements or responses of others while using the instrument, potentially influencing their own usage patterns and responses to the scale. Moreover, despite instructions to refrain from discussing the survey answers among themselves, the nature of the experiment made it challenging to prevent this from occurring. This uncontrolled communication could add an unaccounted-for variable that may impact participants' perceptions and responses.

4.8 Conclusion

In this chapter, I investigated the perceptions of conversation success among PwNH and PwIH, in face-to-face group conversations. Initially, I explored the link between participants' in-the-moment perceptions of conversation success and their overall perception of conversation success reported after the conversations. I found that their in-the-moment assessments are strongly correlated to the assessment after the conversation. Then, I found that when the seven themes of conversation success outlined in Chapter 2 were translated into a survey format, they showed correlations with the overall perception of conversation success, consolidating into a single factor. Finally, this chapter explored the influence of adversity manipulated through variations in background noise levels and the use of hearing aids, on participants' perceptions of conversation success. Results show a decrease in the perception of conversation success in a high level of background noise (72 dBA) regardless of the hearing group or hearing aid usage. No differences in conversation success ratings were found between low and medium background noise levels (30 dBA and 54 dBA). Hearing aids provided a benefit for PwIH

in low and medium background noise levels, making them more similar to PwNH. However, no benefit has been observed in loud background noise levels. But notably, at 72 dBA, PwNH were experimenting similar difficulties as PwIH regardless of the aiding status.

These results provide a foundation for the subsequent analysis presented in Chapter 5, where we will further examine the relationship between participants' perceptions and specific communication behaviours.

Chapter 5
Communication behaviour and success (Study 2)

Below is a list of conferences where parts of this chapter have been presented either in short papers or posters. While some of the figures are reused in this chapter, most of the content presented here is original.

- Poster, ICAIR23: Nicoras, R., Fischer, R., Hadley, L.V., Smeds, K., Godfrey, M., Buck, B., & Naylor, G. (2023) Group conversations: communication behaviour and success, International Conference on Aging, Innovation and Rehabilitation (ICAIR, 2023), May 8th, Toronto, Canada.
- Poster, ISAAR workshop, 2023: Nicoras, R., Fischer, R., Hadley, L.V., Smeds, K., Godfrey, M., Buck, B., & Naylor, G. (2023) Conversation success in small groups, ISAAR satellite workshop on communication, August, Copenhagen, Denmark
- Forum Acusticum conference proceedings: Nicoras R., Hadley LV., Smeds K., Fischer R., Godfrey M., Buck B., Naylor G. (2023), 'Small group conversations: communication behaviour and success', Proceedings of the Forum Acusticum, Torino, Italy, Sep 12th, 2023, A10-02

Matthew Godfrey contributed to data cleaning, pre-processing, and computed derivates of some metrics (e.g.: head orientation to the speaker) analysed in this Chapter. The analyses were conducted by me. Statistical consultation was provided by Oliver Zobay.

5.1 Chapter Summary

In face-to-face interactions, behaviours play an essential role. Both verbal expressions and nonverbal cues contribute to the conversation's dynamics and could influence how participants perceive conversation success. The conversational dynamics are amplified in group discussions, where interlocutors' behaviours must adapt not only to the acoustical environment but

to multiple participants simultaneously. This chapter investigates the behaviours that interlocutors tend to show when conversations are going well, whether they drive success or simply signal its attainment. Conversation behaviours are multimodal (verbal and non verbal), therefore I focus on interlocutors vocal activity and head movement. The findings show that for PwIH talking more relates to success, and as a group, silence is negatively related to success. In addition, individually nodding relates to success, and as a group, looking together towards the speaker is associated with increased success. More details on specific interactions between behaviours and background noise levels, hearing status and hearing aids are reported in the results section of this chapter.

5.2 Background

5.2.1 Communication behaviours

Face-to-face interactions can be demanding at multiple levels, from cognitive processes like attention to observable behaviours. Although often studied in isolation, researchers interested in conversation recognise the importance of linking these processes. Hadley et al. (2022) proposed a categorisation of communication behaviours based on three levels: first, by grouping behaviours according to the modalities or effectors through which they manifest, addressing 'what' people do; second, by organizing behaviours based on their underlying cognitive processes, addressing 'how' these behaviours are generated; and third, by categorizing behaviours according to their social meaning, addressing 'why' people use them.

In this chapter, emphasis is placed on the link between conversation behaviours and success. Here, I explore the 'what' aspect of behaviours, given their relatively straightforward observability and measurability compared to the inner workings of cognitive processes or the subtlety of social meanings. Observable actions, such as speech patterns, hand gestures, and body language, are readily detectable by interlocutors and can be quantified in laboratory settings. These behaviours can be captured and analysed using standardized measurement tools and their objectivity can enhance the reliability and replicability of research findings. Also, these behaviours are often directly

observable by both researchers and practitioners without requiring specialized training or equipment, making it easier to translate research findings into practical interventions aimed at improving communication outcomes in everyday situations. For example, interventions designed to enhance communication skills or reduce misunderstandings can be based on specific observable behaviours identified through research, offering tangible strategies for individuals to implement in their interactions. Furthermore, advancements in technology have enabled the development of devices aimed at monitoring and interpreting these observable behaviours, underscoring the feasibility of measuring and analysing conversation behaviours.

Nevertheless, focusing on 'what' people do during in-person conversations presents a challenge due to the multimodal nature of communication behaviours. Behaviours seldom occur in isolation; rather, they often intertwine across various modalities. For instance, speech is commonly accompanied by gaze, gestures, head movements, and posture adjustments. While these behaviours typically align to enhance the message (congruency), such as maintaining eye contact while talking, there are instances where they may be incongruent, as seen when nodding while verbally expressing "no."

In this chapter, I focus on two primary modalities of behaviour: vocal activity and head movement. While the main body of this chapter concentrates on observed behaviours, the discussion section opens avenues for deeper exploration, occasionally probing into the potential 'why' behind interlocutors' actions. It is worth mentioning that considering a behaviour’s social interpretation (i.e.: what can be inferred from what people do) or a person’s intention (i.e.: what they want to express with their behaviour), can bring some challenges. For example, a simple gesture like crossing arms can convey different meanings depending on the context or the individual. In one setting, crossed arms might indicate defensiveness or disagreement, while in another, it could signal comfort or relaxation. Conversely, different behaviours, such as a nod of the head or a smile, can convey similar messages of agreement or friendliness, emphasising the complexity of interpreting nonverbal cues in communication.

The relationship between communication behaviours and conversation success has the potential to be a two-way street. While conversation success may influence the presence (or absence) of certain behaviours, individuals can also intentionally use specific behaviours to enhance success in conversations. This link is not necessarily causal as conversation success may serve as the context in which certain behaviours manifest, rather than being the direct trigger or motivation.

In group conversations, the perception of success can vary dramatically between individuals and the group as a whole. For instance, in a group of four people, one individual might dominate the conversation and feel highly engaged, while the others may be passive or even bored. This disparity shows how individual perceptions of conversation success can differ greatly from the overall group's assessment. Additionally, conversation behaviours can manifest at both the individual level, such as talking, and the collective level, such as the distribution of speech within the group. Therefore, in this study, I explore the relationship between individual perceptions of conversation success and corresponding behaviours, as well as the association between group perceptions of conversation success and group behaviours.

In the upcoming sections, I examine the roles of specific behaviours in conversation and their potential link to conversation success. First, I discuss three vocal activities commonly observed in conversation: speech, verbal backchannels, and laughter. Next, I look at specific head movements, such as nodding and orienting towards others, within conversations.

5.2.1.1 Vocal activity in conversation

5.2.1.1.1 Speech

Given the central role of speech in conversation, many studies focus on language as a main tool for understanding conversation. In hearing sciences for example, it has been shown how, during conversations, people employ different linguistic strategies as mechanisms to cope with hearing impairment, such as interrogative constructions, clarifications, and repetitions (Lind et al.,

2004, 2010). However, this chapter focuses on the non-linguistic characteristics of speech.

Speech level, a key paralinguistic variable, serves as a tool to understand how individuals adapt their speech in adverse conditions. Lombard speech, a well-documented non-linguistic vocal effect, refers to the phenomenon where talkers adjust their speech production in response to noisy environments. Studies have shown that in adverse communication conditions, talkers tend to produce speech with greater vocal intensity, higher fundamental and formant frequencies, and greater emphasis in specific frequency bands (e.g.: 1-3 kHz range) (Aubanel et al., 2011, 2012; Hazan & Baker, 2011). This effect seems to be enhanced when communicating to other people. Garnier et al. (2010) showed that a greater Lombard effect is seen when individuals were told to communicate with someone else compared to reading in to a microphone. Moreover, speakers adjust their speech to their listener's needs (Beechey et al., 2020; Hazan & Baker, 2011). For instance, normal hearing participants increase their vocal effort in proportion to the level of hearing impairment of their communication partner (Beechey et al., 2020), even when they are not explicitly aware of the hearing loss.

The amount of time spent talking during conversation is another relevant non-linguistic aspect of speech that could relate to conversation success. People who talk more are perceived as extroverts, assertive, less shy (Richmond et al., 2008) and tend to be liked by others more (Hirschi et al., 2023). When it comes to participants with impaired hearing, there is evidence showing that, when listening becomes too difficult, they dominate the conversation as a coping strategy (Wilson et al., 1998). The more they talk, the less they need to listen. Speech production is also impacted by the surrounding environment. For instance, research suggests that when dyads converse in loud background noise, individual speech proportion decreases (Hadley et al., 2021b).

In conversation, interlocutors must coordinate their speech turns, a process known as turn-taking behaviours. These are usually measured by determining the moment when an interlocutor initiates a turn and are known as

the timing of speaker onsets and offsets. Turn coordination refers to the exchange of speaking turns among interlocutors as they manage the progression of the conversation. In linguistic terms, a segment of spoken language is often termed an 'utterance'. When a speaker pauses during their speech, followed by another utterance by the same speaker, this sequence is referred to as an 'intra-speaker gap'. This term specifically denotes the pause interval within the speech of a single individual. An 'inter-speaker gap' refers to the pause between the speech of two different speakers within a conversation. It represents the break in dialogue as one speaker finishes their turn and before the next speaker begins. The 'floor transfer' denotes the temporal transition between speakers during a conversation, indicating the transfer of conversational turn from one speaker to another. The floor transfer offset represents the value that specifies the duration of this transfer: negative if the onset occurs before the previous speaker finishes, and positive if it happens afterwards. Several studies provided evidence on how turn coordination is affected by difficult contexts such as background noise and hearing status: interlocutors with impaired hearing seem to be slower and have more variable floor-transfer offset values and longer utterance durations (Petersen et al., 2022; Sørensen et al., 2021). In terms of background noise conditions, research indicates that floor-transfer offset values are not as precise as in quiet (Heldner & Edlund, 2010). While some studies show that the inter-speaker pauses tend to decrease at higher background noise levels (Hadley et al., 2019), other studies have found the contrary, showing that inter-speaker pauses get longer under similar conditions (Sørensen et al., 2021). However, we do not know what ideal turn-taking dynamics are, what the optimal floor transfer offset is, what the best duration for pause is, nor whether such a thing even exists.

While there is no clear evidence demonstrating how speech relates to conversation success, there are reasons to believe that speech level, speech time, and balanced distribution of speech within the group could play important roles in achieving success. Two themes from Study 1 (Chapter 2) center around speech. Firstly, the theme *Being spoken to in a helpful way* included indications that conversations are successful when speakers use a loud and clear voice

(e.g., statement: 'The speaker has a loud and clear speaking voice.'). This implies a potential relationship between speech level and conversation success. Secondly, the theme *Perceiving a flowing and balanced interaction* indicates that success may be related to individual speech time (e.g., 'Participants are fully engaged and contributing.'), as well as to a balanced distribution of speech among participants (e.g., 'All participants contribute equally.'). Therefore, in this study, we expect that individual conversation success would be associated with individual speech level and speech duration, while group conversation success would be related to balanced distribution of speech within the group.

5.2.1.1.2 Silence

While the average inter-speaker pause lasts around 200 ms (Stivers et al., 2009), the act of preparing even a simple utterance like 'yes' can take up to 600 ms. To maintain the smooth flow of conversation, it is suggested that interlocutors predict their partners' message (Riest et al., 2015) allowing them to prepare a response early and reduce turn-taking gaps. More accurate predictions can therefore lead to shorter pauses. This seems important given that most research suggests that long pauses in conversation carry negative implications. Long pauses are linked to disagreement (Jefferson, 1989; Stivers & Robinson, 2006), poor social skills (McLaughlin & Cody, 1982), feeling uncomfortable (Koudenburg et al., 2011) awkwardness (McLaughlin & Cody, 1982). However, there are some exceptions. Long pauses can be positively perceived in conversations between friends (Templeton et al., 2023) and in doctor-patient communication (Hill et al., 2003) allowing for emotional or/and informational processing. Nevertheless, as our study involves relational group conversations between unacquainted interlocutors, we would expect periods of silence to have a negative relationship with group conversation success. Moreover, periods of silence could impact the perceived flow, one of the themes (*Perceiving flowing and blanaced interaction*) previously identified to drive conversation success (Chapter 2).

5.2.1.1.3 Verbal backchannel

Verbal backchannels are utterances that denote engagement and active listening without being aimed at taking the floor (Schegloff, 1982). Typically, backchannels take the form of brief utterances like 'yeah' or 'mmm', but can also include non-verbal cues like nodding or facial expressions (Niederehe & Duncan, 1974). While backchannelling can serve various functions, it is commonly employed to encourage the speaker to continue. Although they may sometimes be mistaken for affirmative responses, research on their acoustical characteristics has shown that backchannels are usually higher in pitch than other words (Benus et al., 2007). Although a consensus over the average duration of a backchannel does not exist, previous literature suggests that one backchannel can last around 0.5 seconds (Petersen et al., 2023). Backchannels can also be affected by the acoustical environment, with less verbal backchanneling in loud background noise (Petersen et al., 2023).

Verbal backchannels have not been previously studied in relation to conversation success. However, they can indicate engagement, active listening, and encouragement, reflecting the theme of *Being engaged and accepted* (Chapter 2), which may influence how conversation success is perceived. Therefore, in this study, we explore whether an individual's verbal backchannels are related to that individual's perceptions of conversation success.

5.2.1.1.4 Laughter

Laughter is another vocal activity that has been extensively studied in different fields. Acoustically, laughter has been defined as short vowel-like notes (75 ms long on average), with varying fundamental frequency (approx. 502 Hz for females and 276 Hz for males) which occur several times in a regular pattern (Provine, 1996). More recent research shows that involuntary laughter usually lasts longer and has a higher fundamental frequency than voluntary laughter (Bachorowski et al., 2001; Krepsz et al., 2024; Lavan et al., 2016; Vettin & Todt, 2004) Laughter is often accompanied by its visual counterpart, commonly recognized as a smile. While smiles can occur on their

own, laughter seldom happens without a smile accompanying it. However, when people use sarcasm, laughter can appear in the absence of smile. In terms of frequency of occurrence, speech dominates as the most common activity in conversation, followed by smiling and then laughter (Vettin & Todt, 2004). Research indicates that, on average, there are five instances of laughter for every ten-minute conversation (Provine, 1992). Moreover, laughter is found to occur more frequently in the presence of others than when alone watching TV (Provine, 1992), however, it tends to occur more often after speaking than after listening (Provine & Fischer, 1989). Although laughter serves various purposes, it is predominantly associated with expressing positive emotions (Sauter, 2010). In certain contexts, it can also be used to de-escalate negative experiences (Kohler, 2008). Laughter can be mirthful, and polite, but can also show derision or embarrassment (Tanaka & Campbell, 2011). Considered by some ‘the best medicine’, laughter can also serve as a coping mechanism (Glen, 2003). For instance, ‘coping laughter’ can be used to manage disagreements in interactions (Warner-Garcia, 2014). Therefore it remains unclear if laughter in conversation is showing positive or negative emotions, or both depending on the context.

While there is no clear evidence showing that laughter is directly related to conversation success, some indicators suggest its importance. In a study investigating conversation features, good conversation has been associated with humour (Clark et al., 2019), and laughter often signals humour (Sauter, 2010). Additionally, the theme *Feeling positive emotions* (Chapter 2) implies that laughter could be linked to conversation success (e.g., 'Participants are laughing and being funny'). Therefore we expect that individual laughter may be associated with individual conversation success, and we also anticipate that shared laughter contributes to the group's perception of success.

5.2.1.2 Head movement in conversation

In face-to-face settings, head movement plays an important role during conversation, and it usually complements the vocal activity. While listeners typically maintain relatively stable head positions while attending to speakers,

speakers exhibit rhythmic head movements synchronized with their utterances (Hadar et al., 1983). These movements coincide with the peak loudness of speech (Hadar et al., 1983), effectively emphasizing specific linguistic content (Boholm & Allwood, 2010).

Head movements can signal various cues. For example, nodding (oscillating pitch of the head) typically denotes agreement and can serve as a backchannel, encouraging speakers to continue, while head shaking (oscillating yaw of the head) movement indicates disagreement (Jakobson, 1972). However, head shaking by speakers seems to co-occur with speech disfluencies, signalling lexical self-repair processes (Kendon, 2002). Head movement (yaw orientation) can also serve as a turn-yielding cue, signalling the end of a speaker's turn (Maynard, 1987). Even more, the multifunctionality of head movements in conversation includes affect display as studies have demonstrated that head movements enhance emotion recognition during communication (Otsuka & Tsumori, 2020). More head movements seem to be also linked to communication skills (Jensen, 2016; Okada et al., 2016), leadership skills (Beyan et al., 2018), and personality traits such as dominance (Jasen, 2016).

Finally, head movement can function as a compensatory strategy in challenging communication environments, such as during high levels of background noise. Previous studies have demonstrated that individuals may adjust their proximity in response to increased noise and people tend to look more to their interlocutor’s mouth facilitating lipreading (Hadley et al., 2019). Another strategy involves substituting verbal responses with head movements. For instance, in noisy settings like a bustling pub, individuals may nod or shake their heads in response to a question, thereby avoiding the need for vocalization and minimizing the impact of noise on communication effectiveness.

Head movements could relate to conversation success in several ways. Firstly, nodding often signifies engagement in a similar way to verbal backchannels, aligning with the conversation success theme of *Being engaged and accepted* (Chapter 2). This connection is supported by statements such as

'Body language reflecting engagement, e.g., nodding, smiling.' (Study 1, Chapter 2). Secondly, head orientation towards the speaker facilitates eye contact and can improve lip-reading abilities, as noted by participants in Study 1 ('Being able to follow gestures, lip movements, and facial expressions.' and 'Regular eye contact is being made.'). Given that our study involves group conversations, we are particularly interested in moments when interlocutors look together towards the speaker. This behavior may indicate group cohesion, as all listeners are focused on the speaker, potentially contributing to the overall success of the conversation.

5.3 Aims

Previous research has often examined communication behaviours in isolation, overlooking participants' perception of conversation itself and the meaning of these behaviours. Also, most of the previous work involved dyads (Hadley et al., 2019; Petersen et al., 2022; Sørensen et al., 2021) triads (Hadley & Culling, 2022; Hadley et al., 2021b) and human-computer interaction (Otsuka & Tsumori, 2020) overlooking human-to-human group conversations. This chapter aims to investigate communication behaviours in face-to-face human-to-human four-way interactions and link them to the self-perception of conversation success at both individual and group levels. More precisely, this chapter aims to explore what specific behaviours exhibited during group conversations are associated with participants' perception of conversation success.

RQ1 Individual level: What behaviours during group conversation are associated with individuals’ perceptions of conversation success?

Here I relate participants’ individual ratings of conversation success to their behaviour. The aim is to investigate how various behaviours displayed by individuals in group conversations relate to their own perceived success of the interaction.

Based on previous evidence summarised above, we anticipate that conversations will be rated as more successful when participants engage in increased speech, laughter, verbal backchanneling, and head movement. In situations characterized by environmental challenges such as high background

noise or the absence of hearing aids, we hypothesize that individuals will rely more heavily on non-verbal cues, such as nodding.

RQ2 Group level: What collective behaviours are associated with the group level evaluation of conversation success?

Here I relate group level ratings of conversation success (e.g.: the mean of the four participants in a quartet) to collective behaviour measurements (i.e.: that can only be calculated by taking all interlocutors' behaviour into account, e.g., silence time). The aim is to investigate how shared behaviours among group members relate to the group's assessment of conversation success.

On the basis of the previous evidence summarised above, we anticipate that conversations with more balanced speech proportions, less time spent in silence, increased shared simultaneous laughter and greater joint attention towards the speaker will receive higher success ratings. In challenging situations with high background noise as well as conversations without hearing aids, we expect conversational balance to decrease, and the impact of eye contact with the speaker on the perception of group conversation success to become more pronounced.

5.4 Analyses

5.4.1 Variables

5.4.1.1 Conversation success

Conversation success is operationalized by the Factor 1 regression scores derived from the Factor analysis detailed in Chapter 4. Factor 1 itself does not directly represent the ratings participants provided when asked, "How successful would you rate this conversation?", although they are strongly correlated ($r = .89$, $p < .001$). Instead, it encompasses all survey items pertaining to themes of conversation success as outlined in Chapter 2. In addition to being a more robust measurement, using the Factor 1 scores also allows us to use a linear mixed model, a practice that would not be recommended with a 5-point Likert scale.

In individual-level analyses, scores are computed for each participant (n=72) in each conversation (72 participants x 6 conversations), while in group-level analyses, the average scores are computed across the four individuals in the quartet (N=18) in each conversation (18 quartets x 6 conversations). Factor 1 values span from -2 (indicating an unsuccessful conversation) to 1.75 (indicating a highly successful conversation) (Figure 5.1).

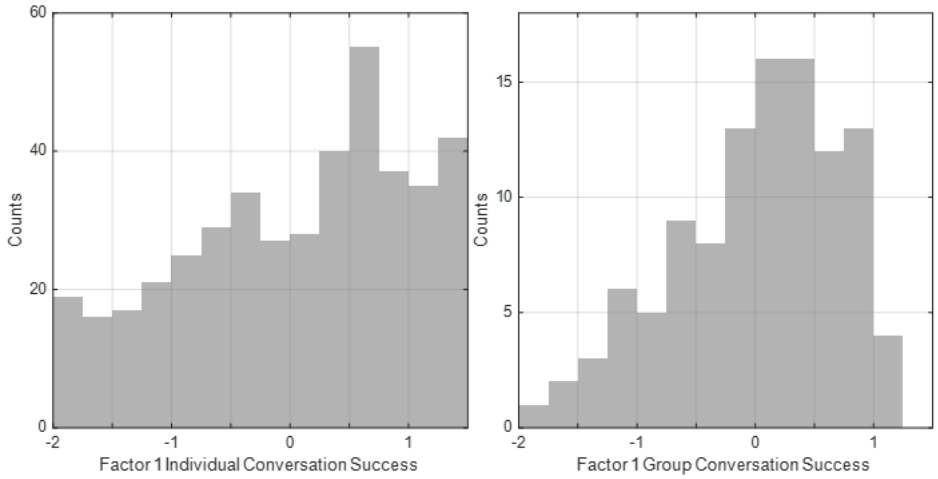


Figure 5.1 Histograms showing the spread of conversation success ratings for Factor 1 in both individual (left) and group (right) level.

5.4.1.2 Communication behaviours

Vocal activity: ELAN 6.5 was employed for voice activity annotation. Voice activity detection was applied on each channel separately (one microphone/person). Initially, an automatic segmentation was performed, splitting uttereances into turns when the gap between them was over 300 ms. This choice was informed by Stivers et al. (2009), who reported a median gap duration ranging from 0 to 300 ms. across multiple languages. Additionally, a dB threshold for detecting speech at all was tailored to each speaker's vocal intensity. The minimum utterance duration was set at 150 ms based on Hadley et al. (2019) reporting short utterances of ~200 ms. This was a conservative way to ensure brief utterances were not missed. Subsequently, manual correction was applied to rectify instances of missed vocal activity or inaccurate detection of non-vocal sounds such as breathing or sighs. The audio recording of each individual was annotated into four classes:

- Speech is defined by any form of voice activity that is greater than 0.5s in duration and involves linguistic content.
- Laughter is defined by vocal activity that follows the laughter pattern of constant frequency peaks. This was manually annotated.
- Backchannel is defined as any form of voice activity that is shorter than 0.5s in duration and not classed as laughter.
- Silence is defined as periods longer than 300 ms without speech, backchannel or laughter.

These four classes are mutually exclusive, meaning that if a person is speaking, the same person cannot express a verbal backchannel and be silent at the same time. While this might not always be the case in real life, as people can laugh while talking, we find this a reasonable compromise due to low occurrence. In such cases, manual annotation prioritised speech over laughter.

The speech material was used to extract the individual time spent talking (s) as well as the individual average speech segment (s)'duration. Also, the speech material was used to derive both the speech level and the signal-to-noise ratio (SNR) for each participant in every conversation. To accurately determine the dB level of participants throughout the conversation, a point of reference is needed. The Harvard sentence spoken by each participant at the beginning of each block was used as a reference (Chapter 3), since we recorded its true dB level with an associated speech waveform. Therefore, we compared the speech level at every point in time during each conversation to this reference and averaged the values for each participant in each conversation. Using the background noise level for each conversation and each participant's position in the room, we derived the SNR. The resulting variables are:

- Speech level: how loud people are talking SPL (dBA) at 1 meter.
- SNR average (dBA): the average SNR received by each participant from all other participants.

From the individual-level classes we can derive a series of group-level classes. From the duration of time each participant spent talking per

conversation, we derive how much imbalance there is in each conversation by computing the difference between the person who talked the most and the person who talked the least. In this way, a higher difference would mean a less balanced conversation while a smaller number would mean a more balanced conversation. Then, we define shared laughter as time in which two or more people laugh at the same time. Similarly, by looking at times when all participants are silent, we extract silence time per conversation per group.

Head movements: translational coordinates (x, y, z) and angular coordinates (yaw, pitch and roll) were recorded at every timestamp during each conversation for each participant (sampled at 100 Hz). From this we derived the participant's distances from one another, and from the centre of the table, as well as their head orientations.

Yaw is the rotation around the vertical axis (shaking left and right), pitch is the rotation around the lateral axis (nodding up and down), and roll is the rotation around the longitudinal axis (tilting left and right). Collectively representing the three rotational movements of one's head.

The yaw values of each participant were transformed so that when they looked straight forward (looking directly at the person opposite the table, assuming nobody moved) the yaw angle was 0°. On average, and from each person's point of view, the other participants were thus situated approximately at the angles 45°, 0°, and -45° (Figure 5.2). If we were to split these angles evenly, the angular separation among interlocutors would be ±22.5° (e.g., if Person A has their head oriented at 20°, we could consider them to be looking at the person in front). However, the default head orientation for interlocutors is towards the person in front. Moreover, when individuals look at someone speaking, they generally do not tilt their heads to face each other directly head-on. Brimijoin et al. (2010) found that the relationship between an auditory stimulus angle and head angle corresponds to a gradient of m = 0.6. Therefore, instead of using the ±22.5° angle delimitation, we used a threshold of ±13.5° (±22.5° x 0.6).

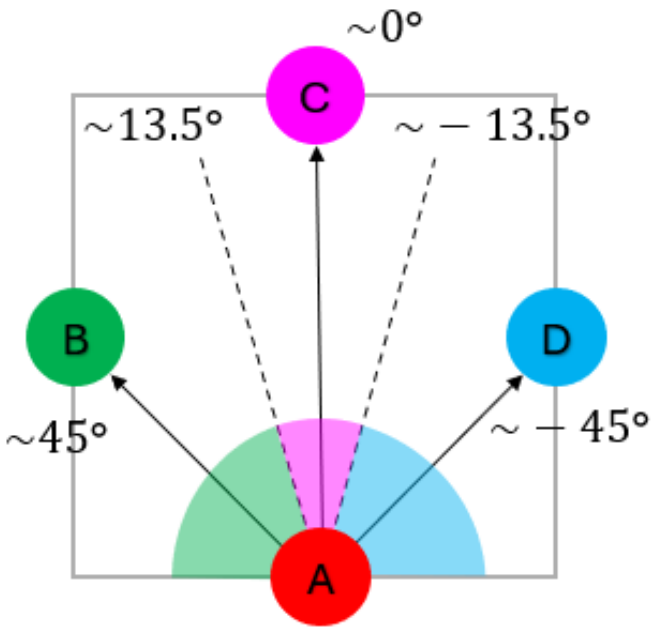


Figure 5.2 Graphic showing the yaw angle distribution from the perspective of participant A.

In the Figure 5.2, the green range represents the degrees in which Person A is considered to be looking at Person B, the pink range corresponds to Person A looking at the person in front, and the blue range indicates Person A's gaze towards Person D.

Similarly, the angular movements in the pitch and roll planes were also adjusted so that when the head orientation is straight ahead, it is considered 0°. The total amount of yaw, pitch, and roll was computed for each person per conversation by summing up the degrees recorded for each person in each plane. For instance, if a person moved their head 20° up and back down, this would count as 40° of movement. These measurements were computed across the whole duration, without accounting for speaking time, backchanneling or laughing.

Combination of head movements and vocal activity

In order to determine intervals during which all listeners were simultaneously looking at the speaker, yaw data and speech data were combined. This resulted in an array containing information for each timestamp and each participant in every conversation, specifying the target person (the

individual to whom the participant is orienting their head) and the speaker (the individual speaking at that moment). This measurement captures a behaviour unique to group dynamics: joint gaze towards the speaker, representing a variable of interest in our analysis.

5.4.1.3 Experimental manipulations

The experiment manipulations involved three variables: background noise level, hearing aid use and hearing status (Chapter 3). These have an impact on the way conversation success is perceived (Chapter 4), and therefore will be incorporated into the statistical analysis as main effects, as well as being examined for interactions with each of the behavioural cues individually.

5.4.1.4 Summary variables

There are various ways to operationalize the same behavioural cue. Table 5.1 provides a concise summary of the communication measures considered in the analysis, along with a brief description of the selected operationalization for each. The selected operationalization is not necessarily the optimal choice but rather a compromise based on the project timeline and the specific objectives of this study. All the variables in Table 5.1 were calculated per conversation.

Table 5.1 Summary of the behavioural cues of interest

Analysis	Modality	Variable (behaviour)	Operationalisation
Individual N=432 18x4x6	Vocal activity	Speech time (s)	Total time (s) spent talking
		Average speech segment (s)	Average length (s) spoken utterances longer than 0.5s
		Laughter time (s)	Total time (s) spent laughing
		Verbal backchannel time (s)	Total time (s) spent backchanneling
		Speech level (dBA)*	The average level (dBA at 1 m.) of each participant's speech
		Average SNR received (dBA)	The average SNR (dBA) received by each participant from all other participants.
	Head rotation	Total yaw movement*	Total amount (degrees) of rotation around the vertical axis (shaking left and right).
		Total pitch movement	Total amount (degrees) of rotations around the lateral axis (nodding up and down).
		Total roll movement*	Total amount (degrees) of rotations around the longitudinal axis (tilting left and right).
Group N=108	Vocal activity	Speech balance	The difference (in percentage of seconds) between the length of speech of the person that spoke the most and the person that spoke the least.
		Silence time	Total amount (s) of time without any vocal activity.
		Shared laughter time	Total amount (s) of time in which two or more participants laughed at the same time.
	Head movement&Vocal activity	Joint gaze towards speaker	The percentage of time in which three listeners have their face oriented towards the same person, and that person is speaking.

*Variables excluded from analysis due to multicollinearity. Details in section 5.4.2

While many other measurements could be derived from vocal activity and head movement data such as speech/backchannel/laughter/backchannel instances, floor holding duration, positive/negative floor transfer offset, average pause duration, proximity, leaning in distance and instances. Although they can provide valuable insights, these remain outside of the focus of the analysis.

5.4.2 Statistical analysis

The lme4 package for R (Bates et al., 2015) was used to build mixed-effects regression models. Post-hoc analyses were done using the ls_means function from the lmerTest package, computing pairwise differences of least-squares means using a Satterthwaite method for estimating the degrees of freedom (Satterthwaite, 1946). Also, the 'emmeans' package for R (Lenth et al., 2018) was used to investigate the estimated adjusted means of the linear mixed-effects models and to compare the interaction slopes. Multicollinearity was assessed using Variance Inflation Factors (VIF) and Pearson correlation coefficients, leading to the identification of several variables that did not meet the criteria. Consequently, to ensure model effectiveness, these variables were removed from the analysis. Specifically, between speech level (dB) and average SNR, a strong negative correlation was observed ($r = -0.63$, $p < .001$), prompting us to retain the latter variable. Similarly, among Pitch, Roll, and Yaw movements, significant correlations were found ($r = 0.58$, $r = 0.63$, $r = 0.59$, all $p < .001$), with the decision made to retain Pitch only for further analysis. These decisions were guided by data visualization indicating a potential larger effect between these variables and conversation success. Both mixed-effects models (individual and group level) followed a stepwise backward approach where variables or interactions were iteratively removed from the model based on statistical criteria using a significance threshold of $p < 0.05$. The next sections provide the initial and the final model for each analysis.

5.4.2.1 Individual level model

The outcome measure used was Factor 1 scores of conversation success from the Factor analysis (Chapter 4). The model included the fixed effects of sound level (30 dBA, 54 dBA and 72 dBA), hearing status (PwIH and PwNH), hearing aids (UA and HA) and the two-way interaction between each of them and the behaviours described above (section 5.3.1.4). No three-way interactions were investigated due to lack of convergence. Furthermore, models included random effects for individuals per quartet number (e.g.: PersonAquartet1).

The initial model: Conversation Success ~ Sound level + Hearing status + Hearing aids + Speech time + Laughter time + Verbal backchannel time + Average speech segment + Average SNR received + Total pitch movement + Sound level*each behaviour + Hearing status*each behaviour + Hearing aids*each behaviour + (1 participant/quartet).

The final model: Conversation Success ~ Sound level + Hearing status + Hearing aids + Speech time + Average SNR received + Pitch + Hearing aids*Pitch + Hearing status*Speech Time + Hearing status* Average SNR received + (1 participant/quartet).

5.4.2.2 Group-level model

For the group-level analysis, the Factor 1 scores (Chapter 4) were averaged across all members of the quartet to get a mean per conversation. The model included the fixed effects of sound level (30 dBA, 54 dBA and 72 dBA) and hearing aids (UA and HA) and the interaction between them and each group-level behaviour (Table 5.1). Hearing status (PwNH and PwIH) was omitted from this analysis due to its balanced distribution within each group and the focus on group-level variables. Furthermore, models included the random effects of the quartet number.

The initial model: Group Conversation Success ~ Sound level + Hearing aids + Silence time + Shared laughter time + Speech balance + Joint gaze towards

speaker + Sound level*each behaviour + Hearing aids* each behaviour + (1 quartet).

The final model: Group Conversation Success ~ Sound level + Hearing aids + Silence time + Speech balance + Joint gaze towards speaker + Sound level*Joint gaze towards speaker + Hearing aids*Speech time + Hearing aids*Speech balance + (1 quartet).

5.5 Results

Below for reference, we provide an overview of the behaviours examined in this study (table 5.2 and 5.3). Although all behaviours were initially considered as potential indicators of perceived conversation success, only a subset were ultimately included in the final mixed models. Subsequently, we present the results of the mixed model analysis, detailing the marginal estimated means and slope contrasts.

5.5.1 Descriptive statistics for communication behaviours

Table 5.2 Descriptive statistics per individual-level behaviour

Behaviour	Aiding Condition	Hearing Status	mean (std)		
			30 dBA	54 dBA	72 dBA
Speech Time (s)	HA	PwNH	84.6 (39.48)	89.51 (35.56)	81.77 (32.83)
		PwIH	74.82 (37.35)	69.9 (34.16)	58.35 (32.24)
	UA	PwNH	94.3 (40.85)	85.49 (41.02)	71.25 (35.45)
		PwIH	57.86 (39.3)	56.46 (38.49)	53.05 (36.41)
Average Speech Segment (s)	HA	PwNH	1.51 (0.56)	1.49 (0.42)	1.54 (0.38)
		PwIH	1.53 (0.44)	1.39 (0.36)	1.4 (0.36)
	UA	PwNH	1.54 (0.42)	1.47 (0.35)	1.4 (0.36)
		PwIH	1.55 (0.5)	1.38 (0.44)	1.39 (0.28)
Verbal Backchannel Time (s)	HA	PwNH	7.13 (3.21)	7.63 (3.36)	7.43 (3.67)
		PwIH	6.23 (3.87)	5.73 (3.66)	5.83 (4.31)
	UA	PwNH	7.88 (3.83)	7.53 (4.46)	8.15 (4.79)
		PwIH	4.01 (3.46)	4.93 (3.49)	5.02 (4.38)
Laughter Time (s)	HA	PwNH	6.38 (7.37)	5.04 (5.13)	4.27 (5.02)
		PwIH	5.11 (5.89)	5.34 (6.37)	5.12 (6.72)
	UA	PwNH	4.9 (4.66)	4.97 (4.86)	4.91 (4.78)
		PwIH	4 (5.18)	5.17 (7.22)	5.16 (8.17)
Average SNR received (dB)	HA	PwNH	24.69 (4.77)	4.27 (4.13)	-5.7 (4.01)
		PwIH	24.51 (4.31)	4.25 (3.5)	-5.89 (3.2)
	UA	PwNH	26.58 (4.27)	5.32 (3.87)	-6.48 (3.71)

		PwIH	25.96 (4.18)	5.3 (3.33)	-6.22 (3.49)
Total pitch movement (degree)	HA	PwNH	2335.75 (877.55)	2465.5 (882.67)	2818.66 (983.61)
		PwIH	2362.32 (831.39)	2530.56 (887.36)	2762.96 (831.79)
	UA	PwNH	2278.49 (1021.85)	2594.79 (785.15)	2778.72 (1179.28)
		PwIH	1972.03 (944.3)	2256.86 (778.34)	2347.03 (903)

Table 5.3 Descriptive statistics per group-level behaviour

Behaviour	**Aiding Condition**	**mean (std)**		
		30 dBA	**54 dBA**	**72 dBA**
Speech balance (%)	HA	0.28 (0.11)	0.26 (0.12)	0.29 (0.13)
	UA	0.35 (0.14)	0.31 (0.13)	0.34 (0.1)
Silence Time (s)	HA	78.04 (15.66)	81.46 (15)	105.34 (23.81)
	UA	70.93 (22.96)	72.76 (17.83)	94.38 (23.57)
Shared Laughter Time (s)	HA	3.24 (3.17)	4.03 (4.26)	3.32 (4.62)
	UA	5.19 (6.26)	3.64 (3.84)	4.05 (4.74)
Joint gaze towards speaker (%)	HA	0.2 (0.11)	0.18 (0.09)	0.14 (0.08)
	UA	0.18 (0.07)	0.17 (0.11)	0.12 (0.06)

5.5.2 Individual-level results

The final model shows substantial explanatory power, with a conditional R^2 of 0.77, indicating that approximately 77% of the variance in perceived conversation success could be accounted for by the included variables. The marginal R^2, which represents the part related to the main effects alone, was 0.47, suggesting that 47% of the variance could be explained by the effects of Sound level, Hearing status, Hearing aids, Speech time, Average SNR received on conversation success ratings.

Consistent with findings from Chapter 4, the main effects of sound level, hearing aids, and hearing status were found to significantly influence the ratings of conversation success. Conversations held at 70 dBA background noise were perceived as less successful than the ones held at 30 dBA and 54 dBA. PwNH tended to rate conversations higher in success compared to PwIH.

In terms of communication behaviours, both speech time and total head pitch movement (degree) demonstrated a positive main effect on the perception of conversation success, indicating that talking for longer and nodding more (pitch head movements) both were associated with an increased level of

conversation success. Furthermore, significant interactions were observed between certain variables.

The interaction between the total pitch movement (degree) and hearing aids (estimate = 0.0198, t = 3.4320, $p < 0.001$) suggests that the effect of pitch on conversation success varies depending on whether individuals use hearing aids. Similarly, the interaction between speech time (s) and hearing status (estimate = -0.0076, t = -3.7210, $p < 0.001$) indicates that the impact of speech duration on conversation success differs based on individuals' hearing status. Finally, the interaction between hearing status and average SNR received (dB) (estimate = 0.0092, t = 2.6310, p = 0.01) implies that the relationship between conversation success and the average SNR received is influenced by individuals' hearing status. These interactions suggest that the effects of pitch, speech time, and average SNR on conversation success are moderated by other factors, such as hearing aid usage and hearing status (Table 5.4).

Table 5.4 Individual-level estimates and p-values for main effects and interactions, with significant results highlighted.

Behaviour	Estimate	Std. Error	df	t value	p-value
(Intercept)	-0.5513	0.3686	224.3145	-1.4960	0.14
Hearing aidsUA	**-0.8606**	**0.1619**	**361.9151**	**-5.3140**	**0.00**
Sound level2	-0.1880	0.2800	202.8730	-0.6720	0.50
Sound level3	**-1.0292**	**0.4171**	**211.9644**	**-2.4680**	**0.01**
Hearing statusPwNH	**0.8545**	**0.1907**	**205.0968**	**4.4820**	**0.00**
Speech time (s)	**0.0084**	**0.0016**	**381.4114**	**5.4040**	**0.00**
Average SNR received (dB)	-0.0031	0.0130	216.1622	-0.2410	0.81
Total pitch movement (degree)	**0.0163**	**0.0062**	**350.9973**	**2.6420**	**0.01**
Hearing aidsUA:Total pitch movement (degree)	**0.0198**	**0.0058**	**327.9967**	**3.4320**	**0.00**
Hearing statusPwNH:Speech time (s)	**-0.0076**	**0.0020**	**401.9854**	**-3.7210**	**0.00**
Hearing statusPwNH:Average SNR received (dB)	**0.0092**	**0.0035**	**280.2662**	**2.6310**	**0.01**

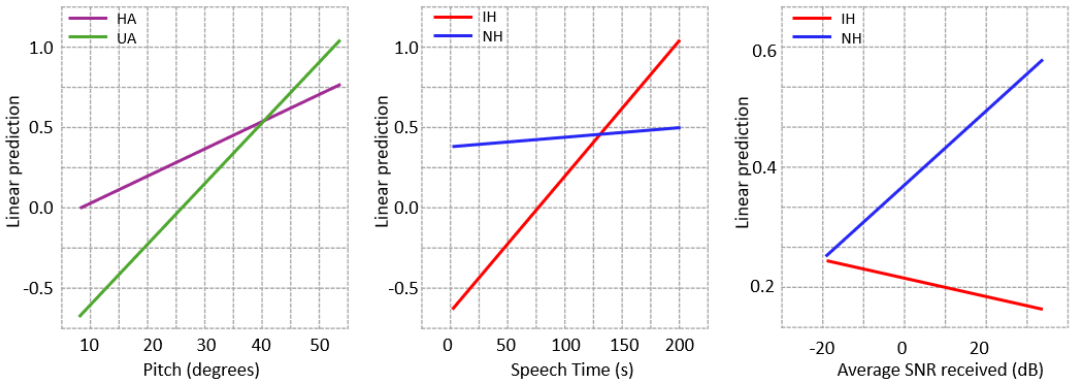


Figure 5.3 Estimated slopes for each interaction in individual level model.

However, estimated marginal means trends (Table 5.5) show that interactions (Figure 5.3) must be carefully interpreted. Regarding the interaction of total pitch movement (degrees) with hearing aids, increasing pitch movement is associated with a higher perception of conversation success in both aided and unaided conditions, with a more pronounced effect observed in the unaided condition. For the interaction between hearing status and speech time, a significant positive effect is observed for PwIH, while no evidence for PwNH. Thus, while our data provides evidence to support that more speech time relates to an increased perception of conversation success for PwIH, we cannot draw this conclusion for PwNH. The interaction between hearing status and average SNR presents unclear findings. Although the results indicate a significant difference between PwNH and PwIH, the specific direction of these effects remains ambiguous for both groups.

Table 5.5 Estimated marginal means for each interaction, with significant results highlighted.

Contrast	condition	trend	SE	df	lower CL	upper CL
Hearing aids:Pitch (degree)	**HA**	**0.0163**	**0.0063**	**366**	**0.00389**	**0.0288**
Hearing aids:Pitch (degree)	**UA**	**0.0361**	**0.0066**	**370**	**0.02303**	**0.0492**
Hearing status:Speech time (s)	**PwIH**	**0.0084**	**0.0015**	**394**	**0.00529**	**0.0115**
Hearing status:Speech time (s)	PwNH	0.0008	0.0014	397	-0.00197	0.0037
Hearing status:Average SNR received (dB)	PwIH	0.0031	0.0134	229	-0.0296	0.0233
Hearing status:Average SNR received (dB)	PwNH	0.0061	0.0132	225	-0.0199	0.0321

5.5.3 Group level results

The model demonstrates a high level of explanatory power, with a conditional R^2 of 0.78, indicating that 78% of the variance in the outcome variable is accounted for by the predictors. Moreover, the fixed effects alone contribute substantially to the model, with a marginal R^2 of 0.67.

Consistent with prior findings, the results reaffirm that both hearing aid use and sound level significantly influence the perception of conversation success, this time at a group level. Regarding group-level behaviours, the analysis shows that silence time has a negative main effect on perceived group-level success, with a coefficient estimate of -0.009 ($t = -3.055$, $p < 0.01$). Conversely, joint gaze towards the speaker is associated with higher ratings of conversation success, indicated by a positive coefficient estimate of 1.173 ($t = 1.718$, $p = 0.09$) (Table 5.6). Additionally, significant interactions were found: the relationship between joint gaze towards the speaker and conversation success varies based on background noise levels, as evidenced by the interaction coefficients for Sound level 2 (54 dBA) (estimate = 1.598, $t = 1.86$, $p = 0.07$) and Sound level 3 (72 dBA) (estimate = 2.460, $t = 2.302$, $p = 0.02$). Moreover, hearing aid usage interacts with speech balance (estimate = -1.927, $t = -3.275$, $p < 0.01$) and silence time (estimate = 0.011, $t = 3.772$, $p < 0.01$). These findings suggest that the relationship between speech balance and conversation success, as well as between silence time and conversation success, differs depending on whether individuals with impaired hearing use hearing aids.

Table 5.6 Group-level estimates and p-values for main effects and interactions, with significant results highlighted.

Behaviour	Estimate	Std. Error	df	t value	p-value
(Intercept)	1.311	0.299	106.168	4.393	0.00
Hearing aidsUA	**-0.680**	**0.291**	**90.933**	**-2.341**	**0.02**
Sound level2	**-0.485**	**0.178**	**92.114**	**-2.731**	**0.01**
Sound level3	**-1.170**	**0.187**	**96.709**	**-6.26**	**0.00**
Silence time (s)	**-0.009**	**0.003**	**106.007**	**-3.055**	**0.00**
Joint gaze towards speaker (%)	**1.173**	**0.683**	**106.889**	**1.718**	**0.09**
Speech balance (%)	-0.681	0.474	100.860	-1.437	0.15
Sound level2:Joint gaze towards speaker (%)	**1.598**	**0.859**	**92.827**	**1.86**	**0.07**
Sound level3:Joint gaze towards speaker (%)	**2.460**	**1.068**	**96.049**	**2.302**	**0.02**
Hearing aidsUA:Speech balance (%)	**-1.927**	**0.588**	**96.604**	**-3.275**	**0.00**
Hearing aidsUA:Silence time (s)	**0.011**	**0.003**	**95.108**	**3.772**	**0.00**

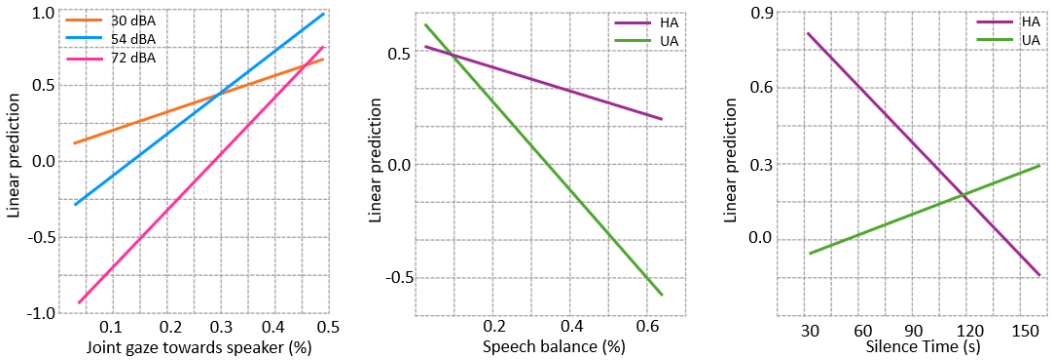


Figure 5.4 Estimated slopes for each interaction in group level model.

As with the analysis conducted at the individual level, it is essential to approach interactions (Figure 5.4, Table 5.7) with caution. For instance, there is insufficient evidence to draw conclusions regarding the impact of joint gaze towards the speaker on conversation success at sound level 30 dBA. However, as background noise increases (levels 54 dBA and 72 dBA), longer durations of joint gaze towards the speaker are associated with an increased perception of conversation success, with a stronger effect observed at 72 dBA compared to 30 dBA (see Table 5.7). While the influence of speech balance on conversation success in aided conditions was inconclusive, evidence suggests that in unaided situations, less balanced speech distribution within the group is associated with a decrease in the group's perception of conversation success. Finally, increased

periods of silence appear to correlate with a decrease in the perceived success of conversations in aided conditions. However, there is insufficient evidence to draw conclusions regarding the association between silence time and conversation success ratings in unaided conditions

Table 5.7 Estimated marginal means for each interaction, with significant results highlighted.

contrast	condition	trend	SE	df	lower CL	upper CL
Sound level:Joint gaze towards speaker	SL1	1.170	0.730	119.00	-0.272	2.620
Sound level:Joint gaze towards speaker	**SL2**	**2.770**	**0.765**	**119.00**	**1.256**	**4.290**
Sound level:Joint gaze towards speaker	**SL3**	**3.630**	**1.053**	**120.00**	**1.548**	**5.720**
Hearing aids:Speech balance	HA	-0.681	0.505	112.00	-1.680	0.319
Hearing aids:Speech balance	**UA**	**-2.608**	**0.490**	**120.00**	**-3.580**	**-1.637**
Hearing aids:Silence time	**HA**	**-0.009**	**0.003**	**118.00**	**-0.015**	**-0.003**
Hearing aids:Silence time	UA	0.003	0.003	116.00	-0.003	0.008

5.6 Discussion

In this chapter, we explored how older adults with and without hearing impairment communicate in group settings. Our main goal was to understand how their behaviours relate to how they perceive the success of their conversations. Participants engaged in group conversations in three background noise levels (30 dBA, 54 dBA and 72 dBA). In half of the conversations, PwIH did not wear their hearing aids, as detailed in Chapter 3.

Our study aimed to answer two key questions: first, what specific behaviours during group conversations affect individuals' own perception of conversation success? Second, what collective behaviours influence how groups evaluate conversation success? We hypothesised that behaviours including speech duration, laughter, verbal backchanneling, and head movement could be most relevant in individual perception of conversation success, while behaviours such as speech balance, silence time and joint gaze towards the speaker would be most relevant in group perception of

conversation success. To achieve these goals, we conducted two separate analyses, which are discussed below.

5.6.1 Individual Conversation Success

The results of the individual-level analysis reveal that both vocal activity and head movements are associated with the perception of conversation success.

Specifically, in terms of vocal activity, our findings indicate that speech time plays an essential role being related to how successful a conversation is perceived by people with impaired hearing. The relationship between speech time and perceived success appears to differ between PwNH and PwIH. Interestingly, PwIH reported significantly higher levels of conversation success when they spoke more. In contrast, the results for PwNH were inconclusive. These findings suggest that speech time may serve as an indicator of conversation success for PwIH but not necessarily for PwNH. Speaking during conversations can facilitate connection, engagement, and a sense of belonging within the group (Guydish & Fox Tree, 2023). Given that individuals with hearing loss often struggle with engaging in group interactions (Bennett et al., 2022), the times that they can actively participate in conversation may potentially enhance their perception of success. It is possible that when PwNH speak less, it's a deliberate choice. In contrast, when PwIH speak less, it is often due to difficulty in hearing. Also, it could be that when PwNH struggle to converse, they may attribute it to external factors, while PwIH might be more likely to internalise the blame (Holman et al., 2023). These findings highlight the importance of facilitating opportunities for individuals with impaired hearing to contribute to group conversations, potentially improving their overall social experiences and well-being.

In terms of head movement, our analysis revealed that nodding up and down during conversations is associated with an increased perception of conversation success. This finding aligns with existing literature suggesting that non-verbal behaviours play a crucial role in enhancing various aspects of communication, including emotion recognition (Otsuka & Tsumori, 2020) and

turn coordination (Maynard, 1987). However, in our study, participants reported higher conversation success ratings when they themselves exhibited this behaviour, not necessarily when their communication partners manifested it. While the rhythmic movement that accompanies speech may contribute to the total pitch movement, the current analysis did not differentiate between nodding while speaking and nodding while listening. Nodding can also be interpreted as a sign of engagement, receptiveness, or agreement with one's conversational partner. Interestingly, we observed an interaction between hearing aid usage and total picth movement on conversation success ratings. While increased total head pitch movement was associated with higher conversation success ratings in both aided and unaided conditions, the effect was significantly more pronounced in the unaided conditions. Conversing without hearing aids, especially for experienced hearing aid users, can be very challenging. Research shows that when facing challenging environments, interlocutors use larger gestures (J. Trujillo et al., 2021). It would not be unexpected if this phenomenon also applies to head movements.

5.6.2 Group Conversation Success

The results from the group-level analysis also show how both vocal activity and head movements are associated with the perception of conversation success.

While in individual-level analysis we demonstrated that talking time is related to conversation success, in group-level analysis it is instead silence time that matters. Silence time seems to relate to a decrease in conversation success regardless of the background noise or hearing aid usage. These findings align with previous studies indicating that long pauses can be perceived as awkward (Templeton et al., 2023) and could trigger negative feelings (Koudenburg et al., 2011). However, it is worth mentioning that the impact of silence time on conversation success ratings may vary depending on the relationship between participants. For instance, when interlocutors are friends or family, the negative effect of silence time might be less pronounced (Templeton et al., 2023). The duration of silence may also be influenced by the conversation topic.

Interestingly, the association between silence time and conversation success appears to be influenced by whether participants are using hearing aids. When individuals with impaired hearing are using hearing aids, prolonged periods of silence relate to a decrease in conversation success for the entire group. However, the results are inconclusive for unaided conditions. One possible explanation is that when aided, there is an expected rhythm to the conversation exchanges, and any deviation from this rhythm may disrupt the flow. Conversely, in unaided conditions, silence time may facilitate conversational flow as interlocutors may require more time to process the speaker's message and formulate responses.

It appears that participants generally prefer more vocal activity during conversations. Not only does reduced silence (presumably accompanied by increased speech) relate to perceived conversation success, but so does speech balance. In group settings, participants acknowledge the importance of balanced interaction, especially in scenarios where hearing aids are not used by the participants who need them. A greater difference between the person who talked the most and the person who talked the least relates to a significant drop in conversation success in unaided conditions. This underscores the negative impact of dominating conversations or excluding certain individuals from participation on the overall success of the conversation.

Another collective behaviour that seems to be associated with the group perception of conversation success is the time spent looking together at the speaker (joint gaze towards speaker). This occurs consistently across different conditions, but it becomes significantly more important for conversation success in environments with heightened background noise. Perhaps, when the conversation is easy (in quiet), joint gaze towards the speaker does not add as much benefit as in a noisy situation. When conversation gets challenging due to background noise, facing towards the speaker might signal engagement and positive effort (“It is difficult, but I am trying!”). Facing the speaker, especially in difficult environments, may also enhance intelligibility by granting access to nonverbal cues such as facial expressions, lip movements

and gestures. While facing the speaker, interlocutors can better signal backchannels or the intention to take the floor or ask for repairs.

5.6.3 Strengths and limitations

While the strengths and limitations of the experimental design have been previously addressed (Chapter 3), this section focuses on the strengths and limitations of the analyses presented in this chapter.

One significant strength hides in the richness of the data explored. This study examined vocal activity and head movement simultaneously, providing a more holistic understanding of communication dynamics. Additionally, the study sheds light on how these behaviours change across different contextual challenges, from quiet to noisy environments.

However, our analyses also come with some limitations that warrant consideration. One limitation is the focus on extralinguistic variables only as opposed to conversation analysis, which may hinder a comprehensive understanding of certain behaviours. For instance, our definition of verbal backchannel as any vocalisation shorter than 0.5 seconds, which despite manual checking and excluding occasional utterances of speech and laughter, could impact our results. Another limitation concerns the measurement of pitch head movement. During the study, participants engaged in an additional task involving a screen and a slider (Chapter 3). Although head movement was not necessary to interact with the screen, participants may have inadvertently moved their heads along the pitch axis while engaging with the screen, potentially confounding the interpretation of head movement data.

The analyses in this chapter consider behaviours as potential predictors of conversation success ratings. However, the manipulated variables are limited to three: background noise level, hearing aid use, and hearing status. It could be argued that in reality, conversation success ratings are not directly influenced by behaviours, but rather by the manipulated variables. Instead, communication behaviours may mediate the effects of background noise, hearing aid use, and hearing status on communication success. For example, in a noisy environment, individuals may lean forward to hear their conversation

partners better. In this case, conversation success ratings could be influenced by the level of background noise, but the effect may be mediated by the compensatory behaviour of leaning in. However, future research can build upon the findings presented in this thesis to investigate how specific communication behaviors serve as mediators in different contexts.

5.7 Conclusion

The current study explored the communication behaviours of older adults with and without hearing impairments in group settings and linked it to perceived conversation success. Through individual and group-level analyses, we identified key behaviours related to conversation success. At the individual level, more speech time emerged to be important, particularly for PwIH. Additionally, total pitch movement during conversations was associated with higher success ratings, highlighting the importance of non-verbal cues. At the group level, speech imbalance and silence time were related to a decrease in conversation success. Periods of silence particularly in contexts when PwIH were using hearing aids were associated to a decrease in conversation success. Increased imbalance, especially in unaided conditions, was related to a drop in the collective conversation success, emphasizing the importance of maintaining balanced interaction dynamics. Furthermore, joint gaze towards the speaker was found to relate to increased conversation success, particularly in challenging environments with increased background noise, potentially due to facilitating comprehension and engagement. Overall, the outcome of these analyses shows how certain behaviours are associated with the perception of conversation success, at both individual and group levels. Hopefully, these results could inform interventions to support older adults in social interactions and promote social inclusion.

Interpreting communication behaviours can be challenging in the absence of verbal content. In the upcoming chapter, I look at how important it is to have access to verbal context and content for evaluating conversation success and interpreting observable behaviours.

Chapter 6
Third-party observations of group conversations (Study 3)

The content presented in this chapter is original and has not been published elsewhere.

This study is an extended version of two BMedSci student projects which I co-supervised together with Dr. Karolina Smeds. Data from these projects are not included here.

Acknowledgements: Andrew Lavens, systems administrator, implemented the survey on the NOTE platform. Dr. Karolina Smeds was involved throughout the qualitative analysis presented in this chapter.

6.1 Chapter Summary

In this chapter, we shift our focus from the perspective of active conversation participants to that of third-party observers, exploring their ability to assess conversation success. We investigate whether observers can accurately assess how successful original intrelocutors perceived the conversation to be, examine how their ratings differ based on the sensory information available, and explore the cues observers base their assessments on. Finally, we use investigate if the conceptualisation of conversation success introduced in Chapter 2 is reflected in the observers' cues of conversation success.

Videos from the previous experiment (Chapters 3, 4, 5) selected as examples of successful and unsuccessful four-way conversations were used as stimuli in this study. Observers watched a series of eight videos, only seeing each conversation in one version: auditory-visual or visual-only. They were asked to rate conversation success during and after each video similarly to the

original interlocutors. Additionally, they were asked to describe the cues that informed their assessment.

By addressing the observer's viewpoint, this chapter aims to deepen our understanding of conversation success, providing a more complete view that goes beyond traditional participant-focused analyses. Observers' perspectives would allow us to assess if there are observable signals of success or difficulty during these conversations. These signals could have practical applications in training frequent conversational partners, helping them learn to recognise and respond to cues of success or difficulty, or in devices like hearing aids or smart glasses, which could detect these cues and adjust settings to improve communication in real-time.

6.2 Background

Conversation research often focuses on what interlocutors perceive (Chapter 4) or what they do (Chapter 5). In previous chapters, I examined the concept of conversation success through the lens of participants' experiences. In the first study (Chapter 2), the participants were respondents in a brainstorming session. In the second study (Chapter 3, 4, 5), the participants were active interlocutors in a group conversation. In this chapter, we shift the focus to observers of conversation. This shift is motivated by several key factors that characterize the observer's viewpoint.

At first glance, passive observers and active interlocutors may seem quite similar. Spitzberg (2000) identifies three entities that can be judged in interactions: the actor, the coactor and the interaction (Figure 6.1). Compared to active interlocutors, observers hold a unique perspective. For example, they are not occupied with formulating responses and they can assess the conversation without the bias of personal involvement. Furthermore, from an outsider's viewpoint, observers can watch all participants at once, noticing things that those involved might miss. Plus, observers cannot influence the conversation's trajectory like the active interlocutors can, which keeps their view unaltered by the expectancies of their own actions. That is why, in this study, I look at how observers' assessments of conversation success relate to those of active interlocutors.

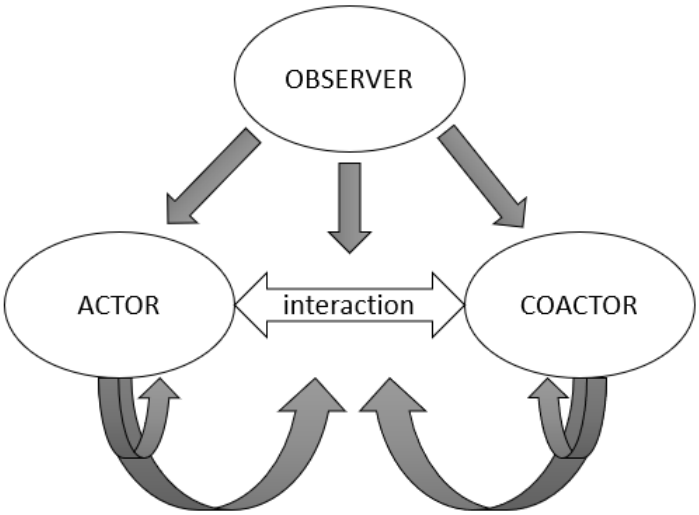


Figure 6.1 Spitzberg's (2000) representation of loci of competence judgements

One common application of an observer's viewpoint in assessing conversations is for use when annotating specific events within the interaction. The observer's perspective has been used to evaluate the presence or absence of specific events (e.g., providing data on whether they detected interlocutors laughing, (Koutsombogera & Vogel, 2022)), as well as to judge different aspects of the overall conversation (e.g., the level of interest of the interlocutors in conversation varying from "not interested in the conversation" to "very interested in the conversation", Gatica-Perez et al, 2005). These events can range from concrete movements and gestures (i.e.: laughter) to more nuanced interpretations, such as interest (Gatica-Perez et al., 2005), involvement (Oertel et al., 2011) and even conversation quality (Raman et al., 2023).

There are reasons to believe that observers may be able to detect if a conversation is successful. For example, research by Ambady & Gray (2002) showed that observers are good at detecting familiarity between interlocutors. In their study, observers viewed silent interactions between opposite-sex dyads and successfully determined the nature of their relationship (strangers, lovers, or friends). This finding indicates that observers can effectively identify the

level of familiarity among interlocutors without actively participating in the conversation. Also, in the current study, both observers and interlocutors share the same age range. This demographic similarity may foster shared perspectives, cultural norms, and communication expectations between observers and active participants. Hence, observers can draw upon their similar experiences to make informed assessments of conversation success, even without direct involvement.

In psychology, this paradigm is recognized as the 'thin slice' approach, where judgments or assessments are formed based on brief observations or limited exposure to stimuli (Ambady et al., 2000; Ambady & Gray, 2002). In the context of interpersonal interactions, thin slicing allows researchers to extract meaningful information from observers watching short snippets of interaction. The assessment can be performed by either trained external observers (Kagan et al., 2004) or non-professional, naive observers (Kumano et al., 2017). However, regardless of their training, observers typically lack prior exposure to the 'thin slice' and possess minimal or no knowledge about the context in which the interaction unfolds (Ambady et al., 2000). While the length of exposure might affect the observer's perception, there is no conclusive evidence supporting a specific exposure length threshold, as durations in prior research can range from as short as 30 seconds (Ambady et al., 2000) to 15-minute segments (Murphy & Hall, 2021). In this study, I used 2-minute snippets of video recordings of conversations from Study 2. Two minutes were judged as long enough for observers to assess whether a conversation was successful or not, and short enough to create a sufficient number of trials.

In addition to the duration of exposure, another essential aspect influencing the observer's assessment of conversation success is the modality in which the material is presented. Visual and auditory stimuli of conversation can evoke different perceptions and interpretations. While visual exposure provides cues from body language, facial expressions, and other non-verbal behaviours, auditory exposure captures tone of voice, pitch, and verbal content. Moreover, these modalities could be presented together (auditory-visual) or

separately (visual-only, auditory-only). The concept of conversation success, as perceived through the interlocutors' experience, encompasses both auditory and visual cues (Chapter 2). While some cues originate from non-verbal behaviour, such as eye contact, others arise from verbal content, exemplified by themes like *Sharing information as desired*. Then, in terms of behaviours during face-to-face interactions, the perception of conversation success was associated to vocal cues (e.g. talking time), non-verbal cues (pitch head movement) as well as a combination of these two (e.g.: joint gaze towards the speaker) (Chapter 5). It remains unclear whether observers primarily base their judgments of conversation success on visual or auditory information. Given that the concept of conversation success is reflected in both, I would expect observers to assess conversation success more similarly to interlocutors when they have access to both visual and auditory cues. This is why, is the current study I wanted to break apart visual-only versus auditory-visual cues to success for observers.

To facilitate an examination of both non-verbal and verbal aspects, it is essential to present two distinct conditions to observers: one featuring exclusively visual stimuli and the other integrating auditory-visual components. Prior investigations have explored the effects of auditory-only, visual-only, and combined auditory-visual stimuli on both PwNH and PwIH, with a focus on assessing emotions rather than conversation success (Most et al., 1993; Most & Aviner, 2009). Notably, studies conducted with children and adolescents indicated that participants with normal hearing consistently outperformed those with impaired hearing in emotion recognition across all sensory conditions (Most & Michaelis, 2012). Furthermore, participants with normal hearing demonstrated superior performance in the auditory-visual condition compared to both auditory-only and visual-only modalities. While these studies primarily addressed emotion perception, there exists an opportunity to adopt a similar experimental paradigm to evaluate individuals' perception of conversation success, an area that remains unexplored in the existing literature. Current research has yet to compare how observers assess conversation success in visual versus auditory-visual conditions during group

conversations. Moreover, to my knowledge, no study has compared observers' ratings of conversation success with ratings of active interlocutors.

Conversations are dynamic, with some moments perceived as more successful and others as less unsuccessful. For instance, in challenging environments with high background noise individuals often struggle to follow conversations (Picou, 2022). In such situations, a common coping mechanism involves relying more on non-verbal behaviour to compensate for hearing difficulties (Trujillo et al., 2021). This increase in non-verbal behaviour due to background noise has also been detected by third-party observers (Paluch et al., 2015). However, it remains unknown whether observers would rate conversation success similarly to active interlocutors in both successful and unsuccessful moments.

The current study aims to address this gap through several key objectives. Firstly, it aims to compare observers' ratings of conversation success with those provided by active interlocutors. Secondly, it seeks to compare observers' assessments of conversation success in visual-only (V) versus auditory-visual (AV) videos, as well as in conversations rated as successful (S) versus unsuccessful (U) by the active participants. This study examines what cues observers use to rate success in these conditions and explores how these cues relate to the seven themes of conversation success identified in Chapter 2.

6.3 Aims

In this section, I present the aims of the study, each accompanied by a corresponding research question. The research questions 1, 2 and 3 are later addressed using quantitative analyses, while the research question 4 is addressed using qualitative analyses.

Aim 1: To assess whether there is a difference between the ratings of conversation success of active interlocutors vs passive observers of group conversations.
Research Question (RQ1*): Is it possible to determine if a conversation was assessed successful without being actively engaged in it?* We expect that third-party observers will be able to accurately reproduce the success ratings of those who were actively participating in the conversation.

Aim 2: To assess whether there is a difference in conversation success ratings when passive observers watch auditory-visual (AV) versus visual-only (V) conversations previously rated by the active interlocutors as successful (S) and unsuccessful (U).
Research Question (RQ2*): Is it possible to determine if a conversation is successful when only given the visual input?* We expect that verbal content improves the accuracy in determining whether a conversation is successful.

Aim 3: To explore the cues that guide the assessment of success in conversations from the observer's point of view.
Research Question (RQ3): *Do observers use different cues to determine what makes a conversation successful in auditory-visual versus visual-only conditions?*
Research Question (RQ4): *Do the seven themes of conversation success (Chapter 2) reflect all the cues used by observers when judging the success of other people's conversations?* We expect the cues employed

by observers in evaluating the success of others' conversations to align with the seven themes of Conversation Success (Chapter 2) and that all seven themes will be reflected in the cues identified by third-party observers.

6.4 Methods

6.4.1 Participants (observers)

Observers N=160 were recruited through the Prolific platform (www.prolific.com)

- Adults (Mean age 58.1 years, SD 6.1 years)
- Normal hearing (self-declared)
- Fluent in English (self-declared)
- Balanced gender distribution (81 females, 79 males)

Once they accepted to participate in the study, observers were directed towards the Nottingham online testing and experimental platform (NOTE). They were paid £10 per hour and took, on average, 34 minutes to complete the experiment. All observers gave their online consent for participation in the study. This research received ethical approval from the West of Scotland Research Ethics Committee (18/WS/0007) and the NHS Greater Glasgow and Clyde R&D (GN18EN094).

6.4.2 Procedure

6.4.2.1 Material

Observers watched 8 videos that had been recorded during the previous experiment (Chapter 3). Each video was a 2-minute snippet out of a 6-minute interaction and involved two people with normal hearing and two people with impaired hearing (quartets). The material was created following the steps below:

1) Videos from quartets that consented were chosen based on conversation success rating. The original interlocutors rated conversation success during the conversation (via a slider) and right after each conversation finished (via closed question with Likert scale options from 1 unsuccessful to 5 very

successful) (Chapter 3). The videos were selected based on the survey question, with "successful conversations" having a mean score across all interlocutors of > 3, and "unsuccessful conversations" < 3. Four videos of each type were chosen. Table 6.1 shows the mean of each video that is used in the current experiment.

2) We cropped 2-minute snippets out of the 6-minute conversations for each video. These were the first two minutes that respected the following criteria: a) no interlocutor explicitly disclosed their hearing impairment via verbal content or gestures; b) the first utterance made sense on its own, i.e. no video began in the middle of an utterance. The videos were censored such that observers would not see how interlocutors were moving the slider during conversation (Figure 6.2), to avoid bias.
3) We created two versions for each video: one for the auditory-visual condition and one for visual-only condition. While the successful/unsuccessful label was fixed for each video, the visual-only/auditory-visual label could change. Each observer only saw one version of each video, and an equal number of AV and V trials. The order of the videos was balanced.

Table 6.1 The conversation success ratings of active interlocutors for each video.

Video AV	Video V	Quartet	Original conversation condition	Individual ratings of success A, B, C, D	Mean of conversation success as rated by active interlocutors	Success status label based on active interlocutors' mean ratings
1a	1b	11	Aided, 54 dBA	5, 4, 5, 3	4.25	successful
2a	2b	7	Aided, 54 dBA	5, 2, 5, 4	4	successful
3a	3b	13	Aided, 72 dBA	2, 1, 2, 1	1.5	unsuccessful
4a	4b	12	Unaided, 72 dBA	1, 1, 1, 1	1	unsuccessful
5a	5b	16	Unaided, 72 dBA	1, 2, 2, 3	2	unsuccessful
6a	6b	8	Aided, 72 dBA	1, 1, 3, 2	1.75	unsuccessful
7a	7b	15	Aided, 54 dBA	5, 4, 5, 4	4.5	successful
8a	8b	9	Aided, 54 dBA	4, 4, 5, 5	4.5	successful

The auditory-visual conditions presented to each participant featured audio recorded via a near-the-mouth microphone. The four audio channels were merged into one and synchronised with the video recordings. All conversations (see Table 6.1) originated from real interactions with background noise. Videos rated as successful were derived from conversations with background noise at 54 dBA, while unsuccessful ones came from conversations with background noise at 72 dBA. Observers could hear the background noise as captured by the near-the-mouth microphone; however, we ensured that the participant's voice remained audible despite the noise.

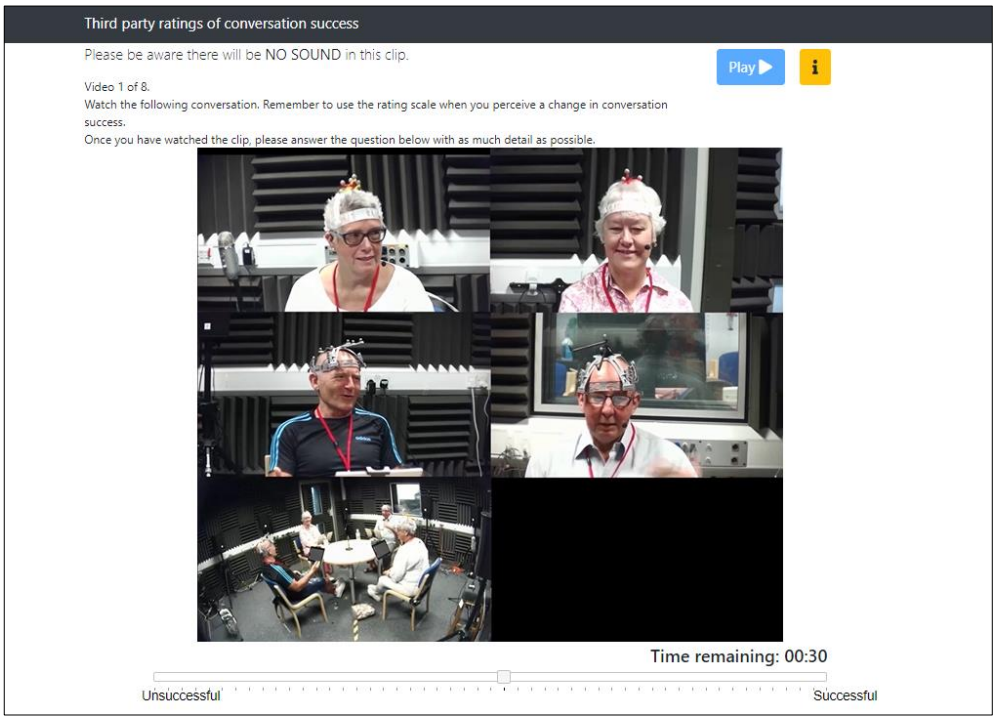


Figure 6.2 Screenshot of NOTE survey showing the observers' view. The slider which observers used to rate conversations can be seen below the video, while the slider which active interlocutors used was blurred for observers.

6.4.2.2 Task

Each observer saw a total of 8 videos (+1 training video). During each video they rated the conversation success using a slider ranging from

"Unsuccessful" on the extreme left to "Successful" on the extreme right (Figure 6.2). Once the video ended, they answered follow-up questions related to the video. The task was structured in two parts:

Part 1: For videos 1-6 observers were prompted with the question "*How would you rate this conversation?*" (Response options are based on a Likert scale ranging from 1 unsuccessful to 5 successful) and two open-ended questions: "*During the video, there may have been moments when you felt the conversation was progressing well. What specific features or actions led you to perceive these as successful instances in the conversation?"* and "*Conversely, there may have been moments when you felt the conversation was not going as well. What features or actions caused you to perceive these as less successful instances in the conversation?".*

Part 2: For videos 7 and 8 observers received the same question "*How would you rate this conversation?*" (Response options ranging from 1 unsuccessful to 5 successful). In addition, a question about identifying the interlocutors with hearing loss was included. Analysis not included here, since the focus of the thesis is conversation success, but may be valuable for future research.

More information about the instructions received by observers can be found in the Appendix D.

6.4.2.3 Design

A two-by-two design was used. The presentation was either auditory-visual (AV) or visual-only (V), and the original conversations were previously rated by the original interlocutors as either successful (S) or unsuccessful (U).

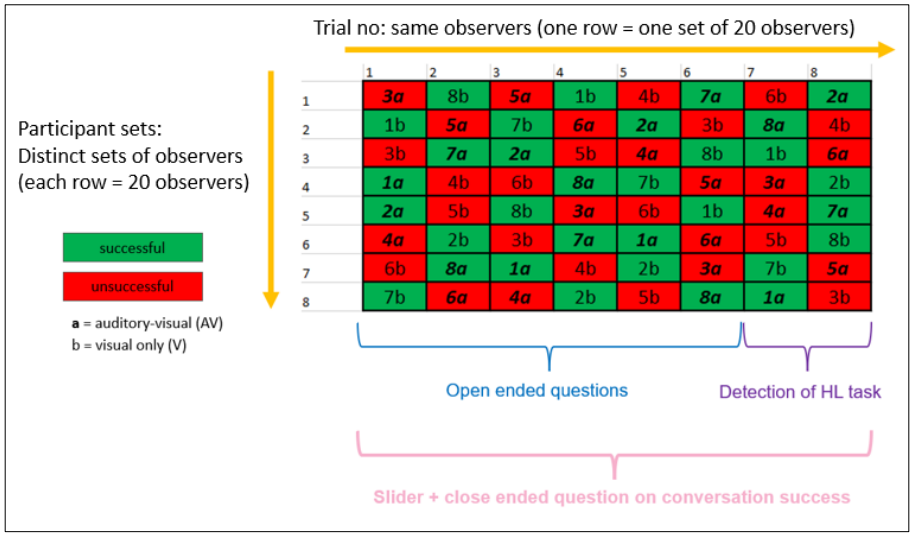


Figure 6.3 Latin square showing the order of conditions presented for each task.

Observers were split into groups of 20. Each group received a different combination of 8 videos (2 videos S-AV, 2 videos S-V, 2 videos U-AV and 2 videos U-V).

Data collected resulted in:

- For all 8 trials participants used the slider and responded to the overall success question, so we collected 20x Slider and overall conversation success ratings for: 16 AV successful videos; 16 V successful videos; 16 AV unsuccessful videos and 16 V unsuccessful videos.
- For 6 out of 8 trials participants gave open answers to explain their success ratings so we collected 20x open ended answers for: 12 AV successful videos; 12 V successful videos; 12 AV unsuccessful videos and 12 V unsuccessful videos.
- For 2 out of 8 trials participants guessed which participants have hearing loss, so we collected 20x detection of hearing loss task answers

for: 4 AV successful videos, 4 V successful videos; 4 AV unsuccessful videos and 4 V unsuccessful videos.

6.4.3 Data analysis

To address Research Question 1 (RQ1), which examines the possibility of assessing conversation success without active participation, we conducted Pearson correlations on the slider data obtained from active interlocutors and from the observers who viewed the interactions. The slider data were matched such that observers' 2-minute slider data corresponded to the same 2-minutes in the slider ratings of the active interlocutors. Each video in each condition was subjected to a separate correlation analysis. Two timeseries (sampled at 10 datapoints/second = 1200 datapoints) means were correlated: the mean from 4 active interlocutors in each video versus the mean of all observers (N=160) who watched the video in a specific condition. This resulted in a total of 16 correlations. Bonferroni correction was applied. Software: R studio.

To address Research Question 2 (RQ2), which examines the potential to recognize conversation success when only given the visual input, we conducted Mann-Whitney and Wilcoxon tests. This analysis used data obtained from the ratings provided in response to the closed-ended questions following each conversation. Specifically, the comparisons were made between V and AV conversations previously categorized by active interlocutors as successful (S) and unsuccessful (U). Software: SPSS.

To address Research Question 3 (RQ3), which investigates whether observers use different cues to determine conversation success in AV versus V conditions, an exploratory content analysis was performed to identify the most recurrent words used by observers in each condition. Word frequency with stemmed words (e.g.: talk, talking, talker were grouped together) for each condition (successful AV, successful V, unsuccessful AV, unsuccessful V) was conducted using NVivo software version 14. The words expressing a behaviour out of top 30 most frequent words are reported in the results section.

To address Research Question 4 (RQ4), which investigates whether the Conversation Success themes outlined in Chapter 2 reflects the cues used by observers when assessing the success of others' conversations, we employed a deductive thematic analysis approach (Braun & Clarke, 2006). This process began with the seven themes of conversation success identified in Chapter 2, which served as the foundational framework. The analysis was conducted using NVivo software version 14. The analysis procedure for RQ4 requires in-depth description.

6.4.3.1 Thematic analysis details

Two researchers (RN and KS) participated in this analysis to ensure coding validity. Both researchers possessed similar knowledge of the seven themes of conversation success described in Chapter 2. The deductive approach involved coding the material with the expectation that it would fit within our pre-existing framework. Observers' responses were segmented into references. A reference represents a set of words or phrases taken from the context that makes sense on its own. Multiple references grouped together by the research team formed a code, and multiple codes formed a theme. Given the deductive approach, the references and codes were placed under pre-existing themes. Note that the existing framework only covers the concept of conversation success. However, observers also provided responses describing 'unsuccessful moments.' To analyse this content, a reverse framework was created. For example, for the theme *Being able to listen easily*, we created a corresponding reversed theme titled *NOT Being able to listen easily*.

A step-by-step methodology for this analysis is outlined below:

1. **Pre-analysis**: Data totalling 34,026 words were exported and arranged into documents corresponding to each condition (S-AV, S-V, U-AV, U-V).
2. **Familiarisation with the material**: This phase involved reading all of the observers' responses.

3. **Co-working**: Two researchers (RN and KS) jointly reviewed 14 responses for each condition (14x4), segmented responses into references, and placed them under appropriate themes in the conversation success framework. References that did not fit specific themes were categorized separately. References fitting multiple themes were coded in several themes.
4. **Separate coding:** RN and KS independently worked on 14 other responses for each condition, then met to discuss and compare the coding.
5. **Automatic coding:** Given the high volume of responses, many of which were duplicated, an auxiliary step was deemed necessary. For each condition, we analysed word frequency using NVivo queries. Words with high frequencies (>10) that clearly expressed cues of conversation success, which RN and KS had previously identified as repetitive, were automatically coded (e.g., "laughter"). A total of 31 different words were automatically coded across all conditions, and their distribution per condition can be seen in Appendix E. The automatic codes were subsequently verified during the manual coding process.
6. **Manual coding:** RN coded the entire dataset while correcting automatic coding. Difficult references were coded separately and discussed with KS.
7. **Co-working (continued):** RN and KS independently worked on another set of responses to ensure agreement remained high. They also decided together how to code difficult references.
8. **Structuring codes and new themes**: KS and RN collaborated on structuring each theme optimally. New themes were created for references not fitting the pre-existing framework.
9. **Extended co-working**: The structure and coding names were reviewed by another researcher (LVH) to validate the hierarchy. LVH, familiar with the pre-existing framework and a native

speaker of English, also ensured appropriate naming of new themes and codes.

10. **Visual representation of results**: Findings are presented in two tables in the results section: one for conversation success cues and another for unsuccessful cues.

Steps 6-9 encompassed an iterative process that involved revisiting previous steps repeatedly to identify subcodes and establish an optimal structural framework.

6.5 Results

Here I first report the quantitative analysis results focusing on similarities between original interlocutors' perceptions versus observers' perceptions of conversation success (Section 6.5.1, 6.5.2, and 6.5.3). Then (Section 6.5.4) I explore the cues used by observers in their conversation success assessment resulting from the qualitative analysis.

6.5.1 Observers versus active interlocutors: ratings of conversation success

RQ1: Is it possible to determine if a conversation is successful without being actively engaged in it?

The correlation results reveal significant associations between the slider data of observers and that of active interlocutors across all videos and conditions (Table 6.2). Correlation coefficients ranged from 0.84 for video 2 in the Visual-only condition to -0.62 for video 4 in the Visual-only condition. Eleven out of sixteen coefficients displayed positive values, showing that observers moved the slider in the same direction as active interlocutors. The rest of the coefficients showed negative correlations meaning that while watching those videos, observers rated conversation success in the opposite direction versus time, compared to active interlocutors.

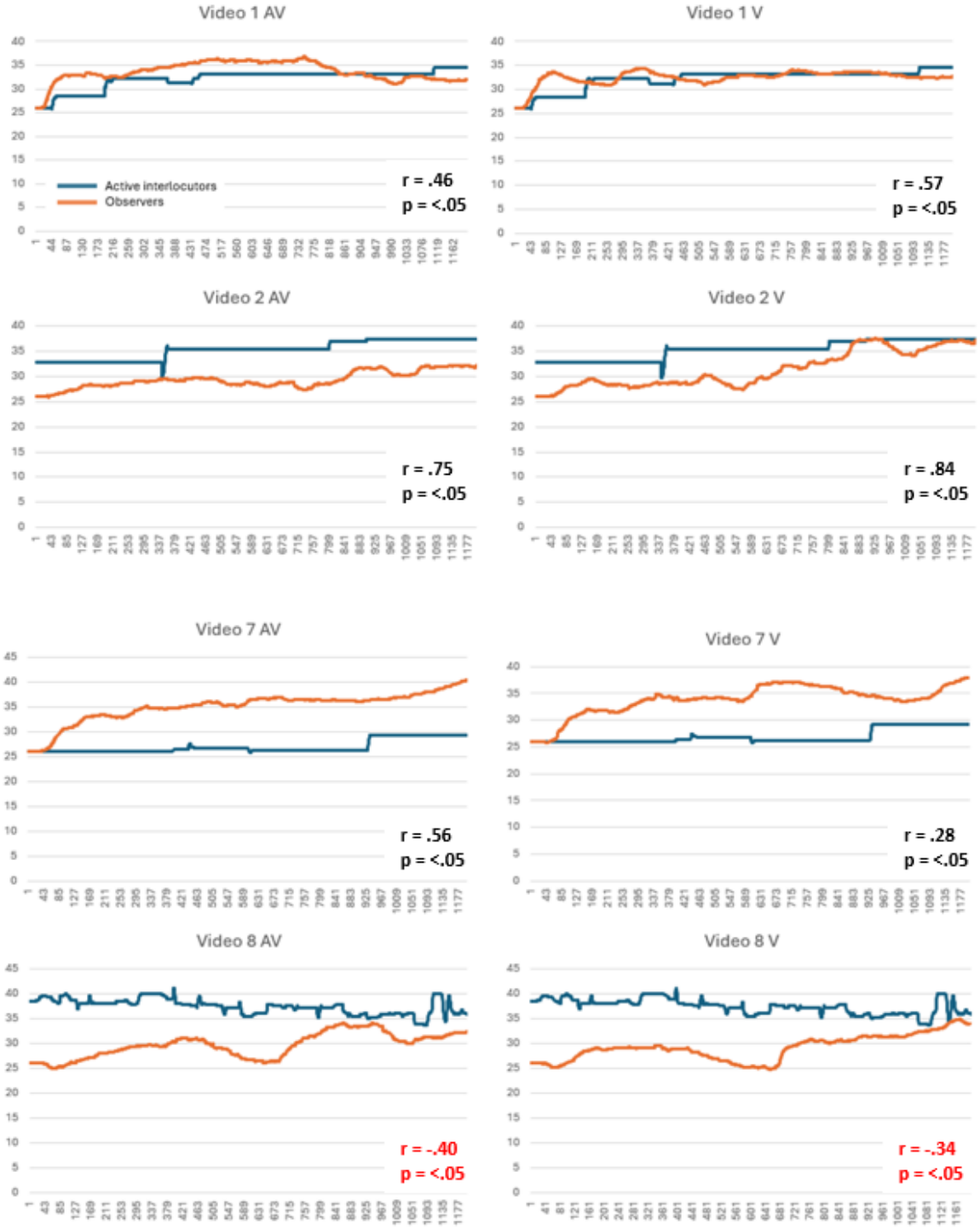


Figure 6.4 Slider data (y axis, 0=unsuccessful and 50= very successful) against time (x axis, 10/s = 1200 datapoints) for successful videos (Videos 1, 2, 7 and 8). Orange represents observers' ratings while blue represents active interlocutors' ratings. Negative correlations are shown in red.

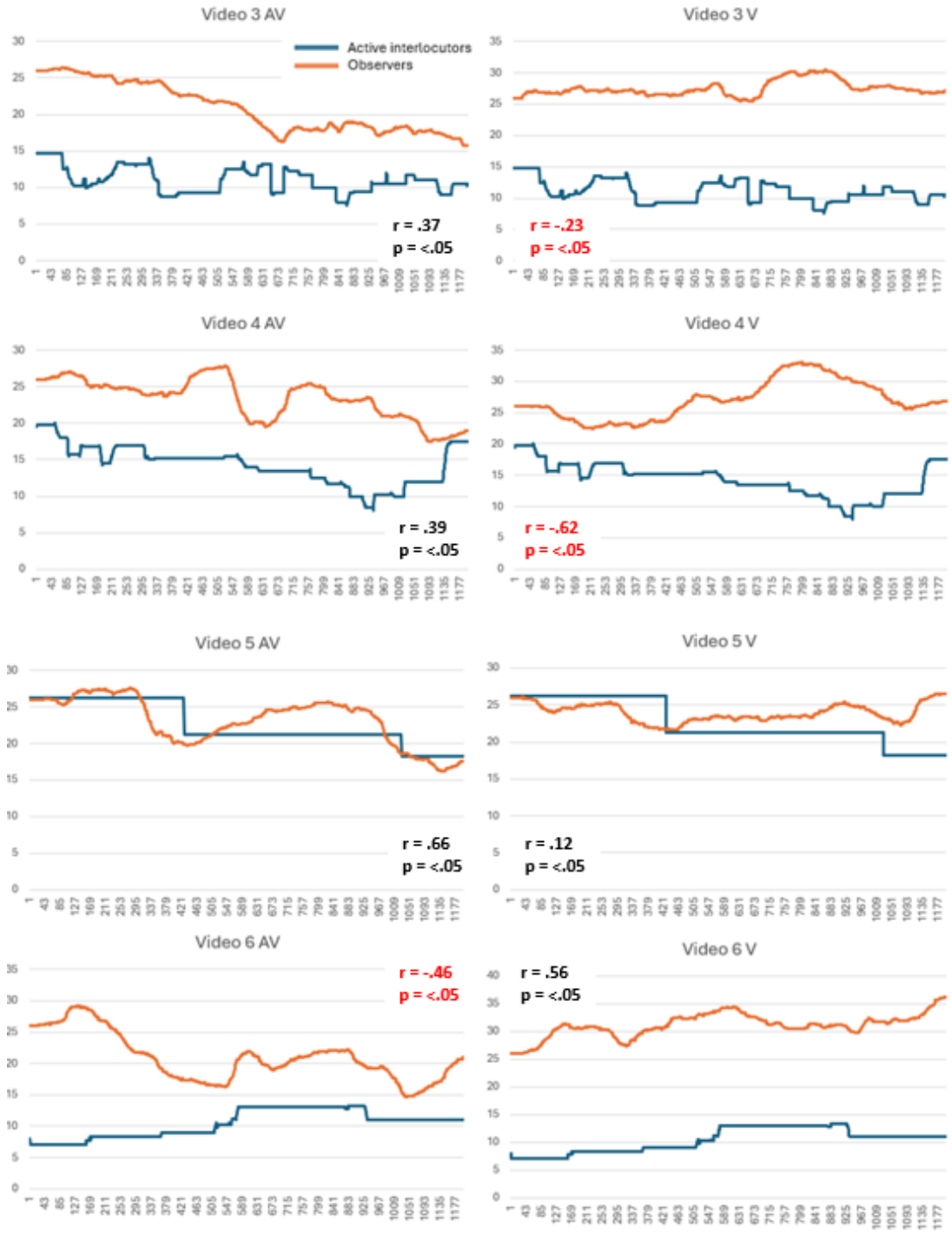


Figure 6.5 Slider data (y axis, 0=unsuccessful and 50= very successful) against time (x axis, 10/s = 1200 datapoints) for unsuccessful videos (Videos 3, 4, 5 and 6). Orange represents observers' ratings while blue represents active interlocutors' ratings. Negative correlations are shown in red.

Results showed that in successful conversations (Videos 1, 2, 7 and 8) observers seem to rate conversation success consistently between Auditory-Visual and Visual-only conditions. For any given successful video, the sign of the correlation in Auditory-Visual condition, is the same as that in Visual-only condition. In Video 8, the correlation sign is negative in both AV and V conditions. However, despite the negative correlation sign, it can be seen (Figure 6) that observers' ratings of conversation success almost matched the ratings of active interlocutors. In unsuccessful conversations the signs of the correlation coefficients of AV and V conditions were in disagreement for Videos 3, 4 and 6. In Video 5, although the AV and V correlations are both positive, the V value is rather small. Taken together, these results suggest that (for reasons yet unknown) observers' assessment of successful conversations are not fundamentally changed by the presence or absence of the auditory signal. With unsuccessful conversations however, their assessment changes dramatically.

Table 6.2 Pearson correlation coefficients for each video in both Auditory-Visual and Visual only conditions.

	Active interlocutors							
Videos	1	2	3	4	5	6	7	8
Observers in AV	0.46*	0.75*	0.37*	0.39*	0.66*	-0.46*	0.56*	-0.40*
Observers in V	0.57*	0.84*	-0.23*	-0.62*	0.12*	0.56*	0.28*	-0.34*

*=p value < .05

The interpretation of these correlations needs to be carefully considered. While correlations are typically computed between independent data points, in our dataset, each data point (slider score) was dependent on the previous one. Additionally, each correlation involves 2x1200 data points, which could lead to overpowering. A binomial sign test was conducted to evaluate whether there was a significant difference between the correlations of AV (auditory-visual) and V (visual-only) conditions. Out of 8 videos, the AV condition had a higher correlation than the V condition in 4 cases, resulting in a proportion of 0.5. The test indicated that this difference was not statistically significant ($p = 1$), with a 95% confidence interval ranging from 0.16 to 0.84.

6.5.2 Auditory-visual versus Visual-only

RQ2: Is it possible to determine if a conversation is successful when only given the visual input?

Wilcoxon signed-rank tests were conducted on participants' ratings after each conversation (each video) for both successful and unsuccessful conversations, revealing significant differences in observers' ratings between the AV and V conditions (see Figure 6.6). Notably, the direction of this difference varied between conditions. In successful conversations, observers rated conversation success higher in the AV condition compared to the V condition ($Z = -2.605$, $p = .009$). Conversely, in unsuccessful conversations, observers rated conversation success higher in the V condition compared to the AV condition ($Z = -6.180$, $p < .001$). These results show that in V conditions observers' ratings of conversation success are closer to the middle of the scale (1 = unsuccessful, 5 = very successful).

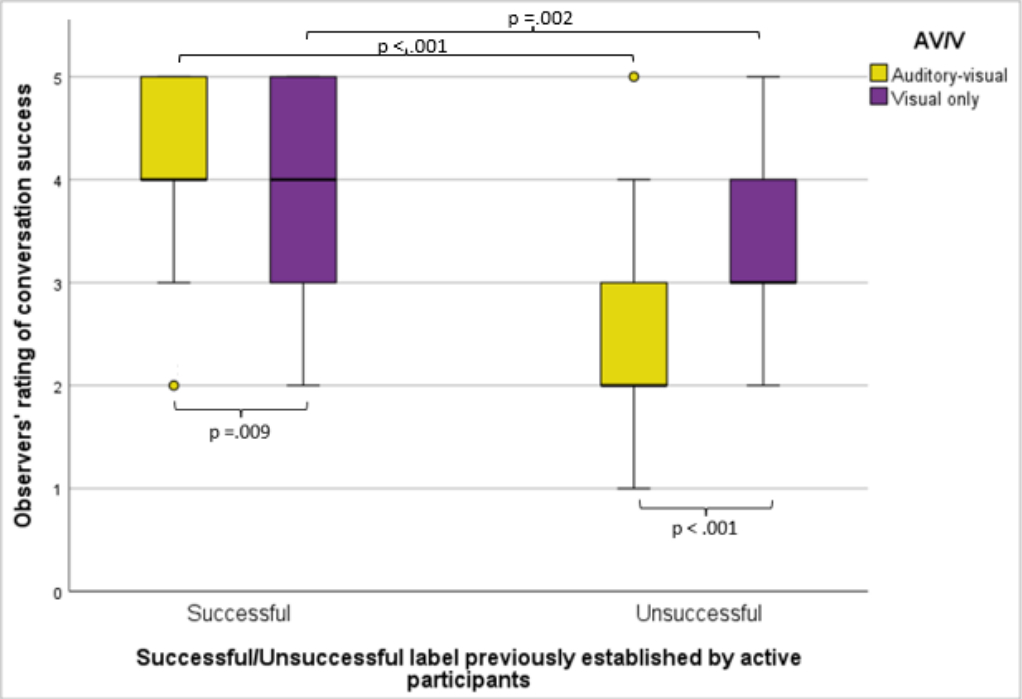


Figure 6.6 Box plots showing the observers' ratings of conversation success for successful/unsuccessful conversations in both AV and V conditions.

Furthermore, Mann-Whitney U tests were employed to compare observers' ratings of conversation success between successful and unsuccessful conversations in both the AV and V conditions. Results indicated that observers rated success significantly lower in unsuccessful conversations compared to successful ones in both conditions (AV: $Z=-9.986$, $p < .001$; V: $Z = -3.143$, $p = .002$). These findings support previous results (RQ1), suggesting that observers can discern whether a conversation is perceived as successful or unsuccessful without direct participation.

6.5.3 Cues of conversation success in auditory-visual versus visual-only conditions

RQ3: Are observers using different cues to determine what makes a successful conversation in auditory-visual versus visual-only conditions?

The exploratory content analysis results show that specific cues are more prevalent in particular conditions (see Table 6.3, full list of words in Appendix F). Observers frequently note non-verbal cues, such as smiling, in V settings, while in AV contexts, they emphasise interpreting verbal content, such as identifying dominance or lack of interest. Moreover, the examination of the top 30 most frequent words indicates that behavioural cues expressed through single words are more common in successful moments compared to unsuccessful ones. This discrepancy may suggest that successful moments are simpler to describe, or that unsuccessful moments require more nuanced descriptions, often necessitating multiple words or negations (e.g., "not/no" + word).

Moreover, it's noteworthy that overall, observers provided more behavioural cues for successful videos in the V condition (2948 references) compared to other conditions: S-AV videos had 1982 references, U-AV had 819 references, and U-V had only 651 references. This observation could imply that behavioural cues of success are more visible than cues of communication unsuccess or that they could require more words due to difficulty in articulating.

Table 6.3 Words that express behavioural cues out of top 30 most frequent words used by observers for each condition and the counts of these words per condition.

Word	Count per condition			
	Successful V	**Successful AV**	**Unsuccessful V**	**Uusuccessful AV**
smiling	189	77	-	-
looks	107	-	151	64
laughing	173	111	-	-
engaging	99	82	39	32
nods	93	30	-	-
gesturing	58	-	-	-
talking	47	43	47	69
joining	42	27	29	-
listening	42	37	-	-
agreement	41	27	-	-
involvement	41	-	-	-
interested	37	40	-	27
speaking	35	-	45	47
trying	34	-	-	-
contribution	-	33	-	-
asking	-	32	-	28
dominating	-	-	40	-
bored	-	-	26	-
expressive	38	-	26	-
struggling	-	-	-	29
hearing	-	-	-	145

6.5.4 Conversation success conceptualisation: observers' perspective

RQ4: Does the Conversation Success map (from Chapter 2) encompass all the cues used by individuals when judging the success of other people's conversations?

6.5.4.1 Successful moments

The framework from Chapter 2 was used for the thematic analysis of the successful moments and the results are presented in Table 3, where the original themes *Being able to listen easily, Being spoken to in a helpful way, Being engaged and accepted, Sharing information as desired, Perceiving flowing and balanced interaction, Feeling positive emotions* and *Not having to*

engage coping mechanisms are shown. Most of the cues that were mentioned by observers fitted into the conversation success map, but one theme remained unrepresented, *Not having to engage coping mechanisms*. Additionally, two new themes emerged: *Engaging effort positively* and *Bonding over difficulty*.

In the study presented in Chapter 2 respondents rated the importance of these themes. The ranking based on their rated importance is not reflected in the volume of cues detected by the observers (Table 6.4). The original respondents considered the most important themes *Being able to listen easily* and *Being spoken to in a helpful way* whereas the most frequently noted behavioural cues from the observer's perspective fell under the themes of "*Being engaged and accepted*" and *Feeling positive emotions*. This observation might reflect the lack of observability of cognitive effort which is reflected in the theme *Being able to listen easily*.

Table 6.4 The themes of conversation success from respondents' perspectives ranked in terms of importance versus the themes from observers' perspective ranked in terms of frequency. The latter reflects how easy the cues are to observe, rather than directly relating to importance.

Respondents' perspective		**Observers' perspective**	
Theme	**Ranking of importance**	**Theme**	**Ranking based on frequency of cues**
Being able to listen easily	1	Being able to listen easily	8
Being spoken to in a helpful way	2	Being spoken to in a helpful way	4
Being engaged and accepted	3	Being engaged and accepted	1
Sharing information as desired	4	Sharing information as desired	6
Perceiving flowing and balanced interaction	5	Perceiving flowing and balanced interaction	3
Feeling positive emotions	6	Feeling positive emotions	2
Not having to engage coping mechanisms	7		
		Engaging effort positively	5
		Bonding over difficulty	7

Regarding the distinction between AV and V conditions, it was observed that most themes were populated by references from both conditions (see Table 6.5). However, there were a few exceptions. References that related to the speaker's voice quality (theme *Being spoken to in a helpful way*) and interlocutors putting effort or bonding over difficulty originated only from the AV condition.

Table 6.5 Observers' cues in themes of conversation success and their corresponding subthemes, source of reference (AV/V), followed by a short description and some representative examples.

Theme	Subtheme	AV/V	Short description	Examples
Being able to listen easily		AV/V	Understanding the speech of conversation partners.	*"They could understand each other over the background noise."*
Being spoken to in a helpful way	Helpful communication partner	AV/V	Collaborative attitude, leading the conversation and inviting other people in.	*"When the person who appeared to be leading the discussion tried to get engagement by specifically asking individuals their opinion or when he looked at them directly if he noticed that they weren't engaging."*
	Helpful voice	AV	Good voice quality: loud, clear, and articulated.	*"Clear and articulate speech from the majority of the group."*
Being engaged and accepted	Positive body language	AV/V	Open body language, hand gestures, mirroring gestures, nodding.	*"The expressive body movements from some members of the group made me think that the conversation was going well."*
	Positive facial expressions	AV/V	Dynamic facial movements, eyes lighting up, and many smiles.	*"Lots of very positive animated facial expressions."*
	Verbal backchannel	AV/V	Verbal feedback such as "yes", "mmhmm".	*"Lots of verbal affirmation on conversation, 'yes' and 'aye'."*
	Engagement	AV/V	Being animated, involved, and listening intently.	*"Everybody engaged in the conversation and debated their preferences and reasonings."*
	Respect	AV/V	Being polite, tolerant, and accepting of different opinions.	*"When they ask thoughtful questions or provide relevant responses that show they're genuinely interested in the topic, it creates a positive atmosphere of mutual respect and understanding."*
	Agreement	AV/V	Agreeing with each other in general or about the topic, being on the same wavelength.	*"When there was agreement about a particular food or drink."*
Sharing information as desired		AV/V	Exchanging ideas and getting their point across.	*"All participants (...) were giving their points well."*
Perceiving flowing and balanced interaction	Flowing interaction	AV/V	Conversation flowing smoothly, without long pauses or overlaps.	*"(...) flowing nicely. All contributing and taking turns to ask questions and opine."*
	Balanced interaction	AV/V	Equal participation and involvement without domination.	*"All participants were able to express themselves without anyone dominating."*

Feeling positive emotions	Enjoyment	AV/V	Feeling happy, having a good time together.	*"Everyone seems [to be] having a good time chatting and smiling and enjoying the conversation."*
	Humour	AV/V	Having fun and laughing together.	*"Some moments where there was good interaction and some laughter and humour displayed."*
	Comfort	AV/V	Being relaxed and friendly.	*"All members seemed relaxed and expressing themselves from the beginning. "*
Not having to engage coping mechanisms				
NEW				
Engaging effort positively		AV	Doing something that usually would not be needed to make the conversation successful.	*"Some valiant efforts to make themselves heard over the noise, at least made an effort."*
Bonding over difficulty		AV	Talking about the difficulty of the situation.	*"When they all talked about the distracting noise, and it felt like a weight lifted from the conversation and they began to interact. "*

6.5.4.2 Unsuccessful moments

Regarding the observers' cues for unsuccessful moments, the reversed framework with seven themes was entirely represented (see Table 6.6). While the majority of themes have references originating from both conditions, there are a few exceptions. Unsurprisingly, the AV condition is the sole source for references related to the speaker's unhelpful voice (theme *Not Being spoken to in a helpful way*), as well as for references related to asking interlocutors to speak up (theme *Having to engage coping mechanisms*). Conversely, the V condition is the exclusive source for references related to leaning in to hear better (theme *Having to engage coping mechanisms*).

Table 6.6 Observers' cues in themes of unsuccessful conversation, codes, condition, description and examples

Unsuccessful moments				
Theme	Subtheme	AV/V	Short description	Examples
Not being able to listen easily		AV&V	Difficulties hearing each other, loud and distracting background noise.	*"They were struggling to hear each other "*
Not being spoken to in a helpful way	Unhelpful communication partner	AV&V	Negative comments about the speaker's attitude and behaviour.	*"Man rocking in his chair seemed to 'boss' the conversation a bit, 'I'm in charge' kind of attitude."*
	Unhelpful voice	AV	Poor voice quality: quiet and low volume.	*"When the quiet lady did speak, she was very quietly spoken and could not easily be heard by the others in the group."*
Not being engaged and accepted	Negative body language	AV&V	Closed body language, folded arms, shaking heads, shoulder shrugging.	*"The younger chap had closed arms - closed body language."*
	Negative facial expressions	AV&V	Facial movements that are interpreted as negative: frowning, not smiling.	*"Faces are grumpy."*
	No eye contact	AV&V	Looking down, looking away.	*"They didn't make eye contact."*
	Lack of engagement	AV&V	No involvement or animation.	*"Not much group engagement."*
	Lack of interest	AV&V	No interest in the topic or conversation.	*"The participants clearly weren't interested in the topic."*
	Disagreement	AV&V	Contradictory opinions, negative answering.	"*No agreement between the 4 people."*
Not sharing information as desired		AV&V	Not getting their point across, creating or feeling confusion, and having blank expressions.	*"A participant who appeared to cause confusion."*
Not perceiving flowing and balanced interaction	Not flowing	AV&V	Three categories were found: Awkward conversation: forced, stilted, lack of cohesion. Long pauses: group giving up, lulls, having nothing to say. Overlap: people talking over each other.	*"The conversation didn't really flow and became stilted."* *"General lack of flow to the conversation. There was an awkward silence about halfway through."* *"People talking over other people."*

	Unequal participation	AV&V	Three categories were found: One person dominating Two persons dominating One person not participating	*"The conversation was being dominated by one person."* *"A lot of the time it was just a conversation between two people rather than the group."* *"One person did not really interact at all."*
Not feeling positive emotions	Lack of enjoyment	AV&V	Boredom, frustration, discomfort, seriousness.	*"There were points where the participants were clearly not enjoying the activity."*
	Feeling excluded	AV&V	One person feeling left out.	*"The older woman looked like she felt left out of it."*
Having to engage coping mechanisms	Leaning in	V	Participants moved their body/head forward interpreted as compensation for hearing difficulties.	*"There were also moments when people were really leaning forward, which I interpreted as not being able to hear sufficiently well."*
	Uncomfortable laughter	AV&V	Nervous or embarrassed laughter.	*"A couple of slightly nervous laughs."*
	Asking partner to speak up	AV	Asking for an increase in voice volume.	*"When the lady at the end said she didn't like pasta nobody could hear her, and she had to be asked to speak up."*
	Asking for repetition	AV&V	Asking a talker to rephrase or reiterate what they said.	*"Think there was once where a speaker was asked to re-iterate what they were saying."*

6.5.4.3 Observers' cues of successful and unsuccessful moments

Here, I describe observers' responses, highlighting cues that go beyond or add nuance to the original conceptualisation of conversation success (Chapter 2).

Some themes encompass a variety of cues from different codes, while others are characterized by a single group of cues that are sufficiently similar to not warrant subdivision into codes. Additionally, observers occasionally used different terminologies to convey similar concepts (e.g., "participation" and "contribution"). Furthermore, the cues can be categorized based on their concreteness. While some cues represent observable behavioural manifestations of interlocutors (i.e.: laughter, hand gesture), others reflect observers' interpretations of these manifestations (e.g., "happiness," "engagement," "agreement").

The theme *Not having to engage coping mechanisms* did not include any cues identified by observers when assessing successful moments. However, when observers assessed unsuccessful moments, they identified several coping mechanisms employed by participants. For instance, when interlocutors were observed leaning in to compensate for their hearing difficulties, observers rated the conversation as less successful. Interestingly, observers also mentioned "laughter" as a cue for unsuccessful moments, specifying that in these instances, the laughter served as a coping mechanism for experiencing negative emotions (e.g., “laughing with embarrassment.”). Asking other interlocutors to speak up or to repeat was also identified as cues indicating unsuccessful moments in conversations.

Engaging effort positively represents a newly identified theme for successful moments derived from observers' recognition of participants' efforts to facilitate the conversation. This theme encompasses references to various behaviours aimed at smoothing the conversation process. Observers' responses cite efforts related to vocal activity (e.g., "Some valiant efforts to make themselves heard over the noise"), helpful attitudes toward others (e.g., "Efforts made to include the person not talking"), and perseverance (e.g.,

"People were trying to persevere with the conversation even though it was challenging").

Bonding over difficulty is another newly identified theme for successful moments arising from third-party observations. It relates to instances where interlocutors discuss challenges they collectively face during conversations, such as high levels of background noise, which leads to the conversation being perceived as more successful by observers. This theme illustrates how, amidst challenging conditions, a conversation can still be deemed successful if the participants establish rapport.

6.6 Discussion

In this chapter, we explored the concept of conversation success from the observers' perspective rather than the interlocutors' point of view. We employed both quantitative and qualitative analyses to explore whether observers can determine if a conversation is successful without actively participating in it. Then, we investigated the impact of verbal content on observers' assessment and sought to identify the cues they used in their evaluations. We analysed these cues through the lens of our pre-existing framework of seven themes of conversation success: *Being able to listen easily, Being spoken to in a helpful way, Being engaged and accepted, Sharing information as desired, Perceiving flowing and balanced interaction, Feeling positive emotions,* and *Not having to engage coping mechanisms.*

We hypothesised that third-party observers would be able to accurately assess the success of group conversations without active participation and that access to verbal content would enhance their accuracy. Furthermore, we anticipated that the cues used by observers in evaluating conversation success align with the seven themes of Conversation Success and that all seven themes would be reflected in the cues identified by third-party observers.

6.6.1 Observers versus active interlocutors: ratings of conversation success

RQ1: Is it possible to determine if a conversation is successful without being actively engaged in it?

The correlation between observers' and interlocutors' continuous (slider) ratings of conversation success was found to be significant overall, with most videos showing positive correlations. This indicates that observers generally rate conversation success similarly to active interlocutors. One possible explanation for this alignment could be their potential ability to empathize with the interlocutors. Observers and the active interlocutors are also similar in age, which may foster feelings of empathy and enable them to resonate with the interlocutors' experiences.

However, some deviations were observed in certain instances, where observers rated the evolution of conversation success in the opposite direction compared to the active interlocutors, resulting in significant negative correlations. In the auditory-visual modality, negative correlations occurred two times (Videos 6 and 8), in the visual-only modality occurred three times (Videos 3, 4, and 8). This suggests that certain aspects of the interlocutors' non-verbal behaviour may have misled observers into perceiving the conversation as more successful/unsuccessful than it was for active interlocutors.

In the case of unsuccessful videos (3, 4, 5, and 6), interlocutores may have perceived the conversations as unsuccessful due to the presence of background noise, creating adverse conditions. These external factors might have disproportionately affected the interlocutors compared to the observers, who could only rely on visual and auditory cues from the video. Although observers in the auditory-visual conditions may have noticed the background noise in videos 3, 4, 5, and 6, they might have still rated the conversations as more successful than the participants did, possibly because they might have seen the interlocutors "Bonding over difficulty." This suggests that observers are seeing the result of the effort that went into making the conversation work rather than being able to judge and weigh the cost of it.

In unsuccessful conversations, observers' ratings seem to change their ratings in Auditory-visual versus Visual-only conditions, while in successful conversations their ratings seem to remain constant. This might suggest that, in unsuccessful conversations, the behavioural cues are harder to interpret than in

successful conversations. One potential explanation could be the consistency of the behavioural cues between visual and auditory visual conditions. For example, interlocutors can use laughter because they find the conversation amusing, but they can also laugh over the difficulty.

6.6.2 Auditory-visual versus Visual-only

RQ2: Is it possible to determine if a conversation is successful when only given the visual input?

Prior research highlights the significance of visual cues in communication (Levinson et al., 2014) and an individual's ability to derive meaning from visual signals alone (Nölle et al., 2021; Trujillo & Holler, 2024). For example, studies have demonstrated that individuals can discern intentions such as rejection or acceptance (Nölle et al., 2021), as well as recognise emotions when presented with muted videos or images (Liu et al., 2022). Moreover, the interpretation of the observed behaviour can change depending on the combination of visual cues available (Trujillo & Holler, 2024).

In the current study, observers' ratings of conversation success confirm that judging a conversation relying only on visual cues might be misleading. Observers are not directly addressed by the speaker, therefore they often judge interactions without prior knowledge about the context or the intention of the interlocutors. When they must rely only on visual cues alone, they might interpret behaviours in ways that diverge from their intended meaning.

In conversations previously deemed successful, observers demonstrate a better recognition of success compared to unsuccessful conversations. However, when deprived of auditory input, they tend to perceive the conversation as less successful compared to conditions where auditory information is available. In the visual-only condition, observers may misinterpret visual cues, leading them to perceive the conversation negatively despite active interlocutors rating it as successful. For instance, if an interlocutor appears to shake their head with a grimace, observers may interpret this as a sign of disgust and disagreement and therefore rate the conversation as unsuccessful. However, when accompanied by audio revealing the

interlocutor's humorous remark (e.g., "*Eww, I don't like cheese!*"), observers may shift their interpretation to take account of the context.

Conversely, in conversations previously rated as unsuccessful, observers tend to rate conversation success higher in the visual-only condition than in the auditory-visual condition. In such instances, observers may once again misinterpret visual cues, perceiving laughter or apparent enjoyment as indicators of successful conversation. However, when paired with audio indicating interlocutors' difficulty in hearing (e.g., "*I cannot hear. Do you find it as hard to hear as I am?*"), the interpretation of visual cues may shift towards a less successful interaction.

These findings confirm that during conversations individuals use conflicting cues to convey their message (Grebelsky-Lichtman, 2021) and indeed are explained by the discrepant verbal-nonverbal profile theory (Grebelsky-Lichtman, 2017). In our study, it appears that verbal content played a significant role in enhancing observers' ability to accurately assess whether a conversation was perceived as successful by its original interlocutors. These findings suggest that the ratings given in the audio-visual condition were more similar to the interlocutors' ratings of conversation success. However, to fully comprehend the importance of integrating visual and auditory information in discrepant communication, an auditory-only condition would be necessary.

6.6.3 Cues of conversation success in auditory-visual versus visual-only conditions

RQ3: Are observers using different cues to determine what makes a successful conversation in auditory-visual versus visual-only conditions?

Examining word frequencies facilitated direct comparisons between visual-only and auditory-visual stimuli, revealing that in the visual conditions, a much larger emphasis was given to the movement behaviours than in the AV conditions. This can be seen especially in successful moments. For example, in the visual-only condition observers mentioned smiles and laughters 189 and 107 times whereas in the auditory-visual condition, these words were mentioned 77 and 71 times. The word "engaged" was ranked first in the auditory-visual condition, while in visual-only it was ranked after smiles and

laughter. When observers had access to auditory-visual content they appeared to interpret interlocutors' behaviours using more abstract concepts such as engagement, agreement, and involvement. However, when limited to visual content only, observers relied more heavily on descriptors of observable behavioural manifestations, such as smiling, laughter, and nodding.

6.6.4 Conversation success conceptualisation: observers' perspective

RQ4: Do the Conversation Success themes (Chapter 2) encompass all the cues used by observers when judging the success of other people's conversations?

This study applied the conceptualisation of conversation success to observers perspectives to assess its relevance and extend it as necessary. Results show that observers' and respondents' perspectives on conversation success seem to align for six out of seven themes.

Unsurprisngly, the theme *Not having to engage coping mechanisms* was not reflected in the cues detected by observers. This absence is understandable because assessing conversation success based on the absence of certain behaviours, particularly undesirable ones, can be counterintuitive. Instead, it is often easier to identify the presence of coping mechanisms in unsuccessful moments, as demonstrated by the reverse theme, *Having to engage coping mechanisms*. Hence, the lack of cues in this theme may indeed confirm that successful conversations involve the avoidance of resorting to coping mechanisms.

6.6.4.1 Conversation success

The concept of conversation success as seen through an observer's perspective includes eight distinct themes: *Being able to listen easily, Being spoken to in a helpful way, Being engaged and accepted, Sharing information as desired, Perceiving flowing and balanced interaction, Feeling positive emotions,* and the newly identified*: Engaging effort positively,* and *Bonding over difficulty.*

From the observer's perspective, *Being able to listen easily* appears as a requirement for a successful conversation rather than a cue. The other cues detected (e.g.: taking turns) were possible due to the fact that interlocutors

seemed to be able to hear each other (e.g.: “The room was quieter so everyone was able to hear the conversation. They used eye contact, facial gestures and took turns with the conversation.”).

From the observer’s perspective, the theme *Being spoken to in a helpful way* refers now to both the role of the listener (being spoken to) and the role of the speaker (speaking to others in a helpful way). Previous research showed that individuals adapt their voice volume not only to the environment but also to their communication partners’ needs (Beechey et al., 2020; Hazan & Baker, 2011). These vocal accomodation behaviours were recognised by observers in this study as indicators of conversation success (example). Interestingly, observers also acknowledge participants' communication management skills. While these skills may not be as evident in one-on-one interactions, observers perceive successful group conversations where one interlocutor takes the lead and actively encourages others to participate by directly addressing them (e.g.: “The conversation was clearly led at times by the man who had a voice loud enough to be heard over the background chatter, and his confidence in carrying on seemed to be somewhat infectious. He also successfully engaged other participants by name, and the quieter people were encouraged into contributing.”).

Being engaged and accepted is the most populated theme (with over 900 references). This might be explained by the fact that it includes a large number of cues expressing behavioural manifestations that are also prominently observable in visual-only conditions. Conversations are perceived as more successful when interlocutors use more hand gestures, open body language, and more head movements such as nodding. Observers’ assessments align with the existing literature suggesting that manual gestures enhance communication efficacy by conveying messages more effectively (Holler & Wilkin, 2011; Trujillo & Holler, 2024). Also, there is evidence that nodding is associated to the self-perception of conversation success (Chapter 5). Nodding often indicates agreement, but it can also serve as a backchannel, encouraging speakers to continue (Jakobson, 1972) which might enhance their engagement and feelings of acceptance. Observers also made use of interlocutors’ verbal

content to judge conversation success. They interpreted verbal backchannels as interest and agreement (e.g.: “Non-verbal body actions such as a person nodding in agreement to another's comment(s), facial expressions along with verbal affirmation.”).

Sharing information as desired was also considered important. The cues associated with this theme reflect the interlocutors’ abilities to convey their thoughts clearly to others. However, since observers were unaware of the interlocutors’ intentions, it was challenging to evaluate whether the information exchange fits the intended purpose. Additionally, the topics of conversation in these videos were more social and less transactional, making it difficult to discern the underlying goal of the interactions. While many conversation paradigms focus on sharing information (Baker & Hazan, 2011; Beechey et al., 2019; Garrod & Pickering, 2004), identifying clear instances of information exchange in natural conversations can be challenging from an observer’s perspective.

The theme *Perceiving flowing and balanced interaction* is represented by cues suggesting that both flow and equal participation are indicators of conversation success. From the observer’s perspective, flow is recognised when interlocutors manage to exchange rapid turns, without overlaps or long pauses. These results confirm existing research suggesting that conversational flow may be negatively affected by delays or pauses during interaction (Koudenburg et al., 2017). Additionally, moments of silence in group conversations have been found to diminish perceptions of conversation success (Chapter 5). However, some studies suggest that longer pauses may be acceptable in conversations between close friends (Templeton, 2020) and in specific cultural contexts (Acheson, 2008). Observers also identify speech balance as a cue for conversation success, aligning with our previous findings (Chapter 5) showing that larger disparities in speaking time between individuals can lead to reduced perceptions of conversation success within groups.

The theme *Feeling positive emotions* affirms that enjoyment, humour, and friendliness are indicative of conversation success. Observers noted that

interlocutors in successful moments appeared to be relaxed, having fun, laughing together and joking with each other. These cues were described using various terms such as 'enjoyment,' 'happiness,' and 'amusement'.

The newly identified themes of *Engaging effort positively* and 'Bonding over difficulty' expand the original conceptualization of conversation success (Chapter 2). These themes may not have been included in the initial conceptualization because they represent behaviors that, while contributing to conversation success, are not necessarily desired from the outset. In an ideal scenario, successful conversations flow effortlessly without the need to manage any difficulty. However, in reality, conversations can be challenging. In such instances, the way interlocutors manage these difficulties can lead to successful moments.

Thus, from the observers' point of view, a conversation can have successful moments when interlocutors are *Engaging effort positively*. Cues in this new theme seem to encompass any action conducted by interlocutors that is interpreted by observers as something that in ideal circumstances would not be needed but could improve communication in the current situation. These cues align with existing literature on communication effort. Beechey (2020) describes communication effort as "the interactive feedback system that interlocutors form together during conversation in order to signal receptive difficulty and efficiently compensate for this difficulty.". The effort could be related to anything that could benefit the conversation: maintaining flow ("made the effort to respond"), participating ("made the effort to get involved"), topic ("tried to come up with ideas"), including others in the conversation ("Efforts made to include the person not talking.") and not giving up ("People were trying to persevere with the conversation even though it was challenging").

Finally, *Bonding over difficulty* represents a new theme that emerged from the perspectives of third-party observers. Interlocutors either used humour to alleviate the discomfort of the situation (e.g., 'They were all laughing about how they couldn't hear each other') or expressed similar reactions (e.g., 'When they agreed about how terrible the noise was; they were all engrossed'). The

shared experience appears to bring the interlocutors together (e.g., 'Bonding over the shared experience of difficult conditions, such as noise'), which led observers to judge these moments as successful in spite of the potentially challenging conversation situation.

6.6.5 Strengths and Limitations

This study's primary strength lies in the quality of the responses provided by the observers. Despite being an online study conducted via the Prolific platform, where observers were not directly supervised, most of them diligently watched the videos and provided thoughtful responses to the open-ended questions. For example, one observer commented on what could have been done to improve the conversation: "*It didn't occur to me at the time, but on reflection no one picked up on her difficulty due to the background noise - so although they tried, unsuccessfully, to include her in the conversation, they didn't pick up on the issue she was having (despite her attempt to communicate it), and didn't try to either adjust their behaviour or discuss ways to overcome it e.g. by speaking louder, slower, or more directly facing her. This effectively resulted in leaving her out of the conversation, even if it was unintentional.*"

Additionally, this study linked observers' ratings to those of the original interlocutors. The videos shown to the observers were collected within our department, streamlining the process. The use of pre-recorded videos is an advantage of this study, as it enables the matching of observers' ratings to those of the original interlocutors. Also, this provides the opportunity for future research to correlate changes in observers' assessments of conversation success with specific recorded behaviors exhibited by interlocutors. Moreover, conducting the qualitative analysis on conversation success cues inductively, starting from a framework created by the authors, is advantageous. However, we are not excluding the possibility that an inductive thematic analysis, conducted without prior knowledge of the seven themes of conversation success, would have been at least equally interesting.

While this study examined both visual-only and auditory-visual conditions, a potential future direction could involve the inclusion of an auditory-only condition. This would provide a more complete analysis by

independently testing and contrasting each sensory modality before exploring a condition where all modalities are combined.

6.7 Conclusion

This study aimed to investigate whether third-party observers could detect if a conversation was successful without actively taking part in it. Moreover, the study set out to establish whether third-party observers rated conversation success differently when presented with visual only compared to auditory-visual stimuli. Finally, we explored the cues used by observers to determine if a conversation was successful or not and used these cues to expand the conversation success conceptualisation presented in Chapter 2.

Our findings indicate that observers' ratings of conversation success generally align with those of the original interlocutors, suggesting that individuals can discern the success of others' conversations without direct involvement. However, when observers rely solely on visual information, they may misinterpret unsuccessful conversations as successful. Reliance on non-verbal information might be misleading, as interlocutors can provide contradictory information such as smiling while commenting that it is a very difficult situation. This may have been even more likely for the interlocutors in the videos we used, as they were unfamiliar with each other and likely on their best behavior.

Conversation success, as seen from the observers' perspective was found to include the following eight themes: *Being able to listen easily*, *Being spoken to in a helpful way*, *Being engaged and accepted, Perceiving flowing and balanced interaction*, *Feeling positive emotions, Engaging effort positively*, and *Bonding over difficulty.* These themes extend the conceptualisation outlined in Chapter 2, advancing our understanding of the multifaceted nature of conversation success.

Chapter 7
Conclusions

7.1 Thesis summary

This thesis has been concerned with operationalising conversation success with a focus on older adults with hearing loss. While everyone aspires to have successful conversations, individuals with hearing loss often report difficulties in engaging effectively. Despite the promises of hearing aids to improve listening experiences, their effectiveness in real-life communication scenarios remains limited. Traditionally, assessments of hearing and hearing aids have been centred on listening and comprehension tasks. However, hearing aids are not just tools for passive listening; they are also expected to facilitate engagement and maintain conversations. Conversations, unlike listening-only tasks, are interactive, spontaneous, and dynamic, involving coordination with at least one other person. Therefore, evaluating conversation success goes beyond assessing listening skills alone. Before attempting to measure conversation success, it is essential to first understand what it means, particularly within the context of older adults with hearing loss.

The studies presented in my thesis aimed to understand the concept of conversation success from the point of view of older adults with both normal and impaired hearing, and to identify behavioural indicators of conversation success during group conversations. The research unfolded in a sequential manner, with each subsequent study building upon the insights gained from its predecessor. Study 1 (Chapter 2: conceptualisation of conversation success) laid the foundation for Study 2 (Chapters 3, 4, and 5: self-reported ratings of conversation success and behaviour in four-way conversations), which built upon its findings. Study 3 (Chapter 6: third-party observations of conversation success) then expanded upon the results and materials gathered from both Study 1 and Study 2. This cascade of connected studies allowed for a deeper

understanding of the different layers involved in perceiving successful conversations.

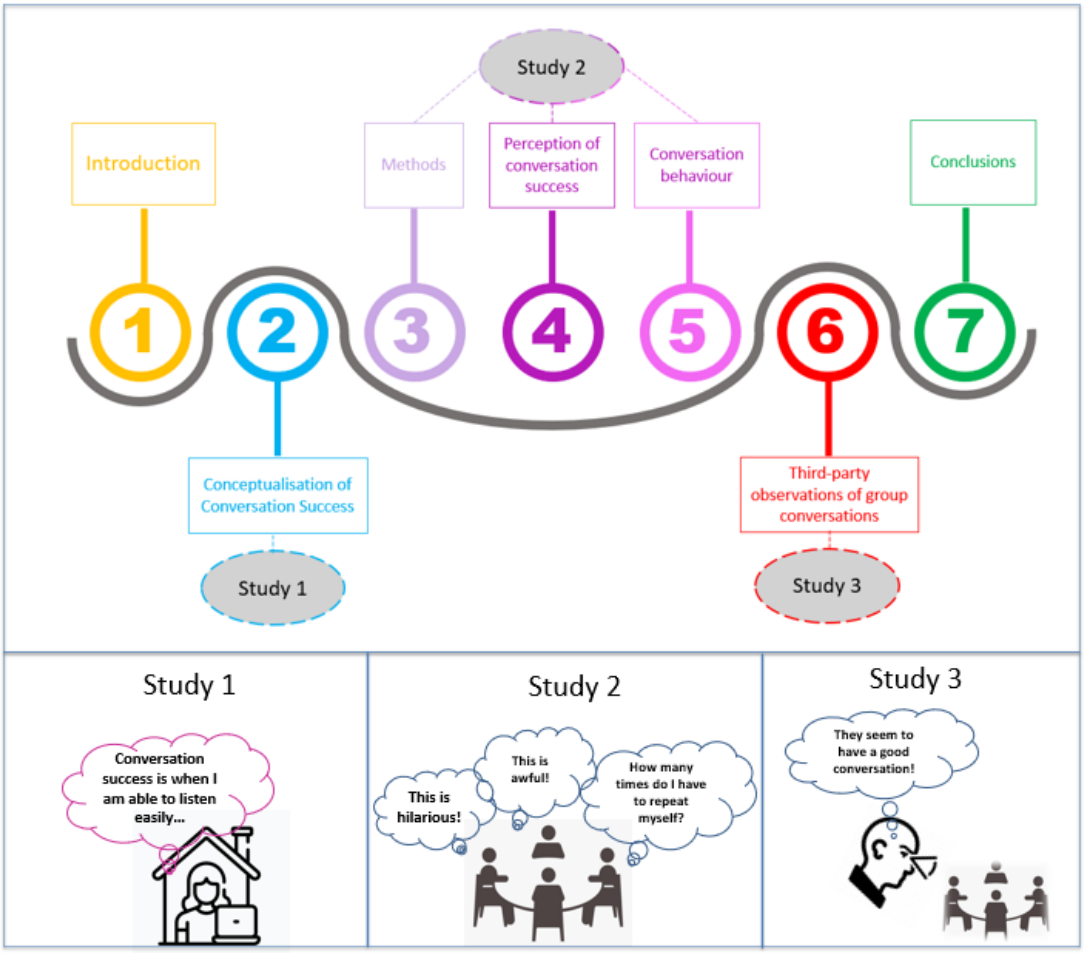


Figure 7.1 Diagram showing the sequence of the studies and their corresponding chapters.

7.2 Major findings

The experiments included in this thesis investigated conversation success from three distinct perspectives: the subjective viewpoint (*what is conversation success?*), the behavioural viewpoint (*what behaviours are associated with the perception of conversation success?*) and the perspective of third-party observers (*how do third-party observers judge conversation success?*).

7.2.1 Subjective perspectives on conversation success

What is conversation success (Chapter 2)? Here I demonstrated that the concept of conversation success is multifaceted. Older adults with normal and

impaired hearing perceive a conversation to be successful when they are *Being able to listen easily, Being spoken to in a helpful way, Being engaged and accepted, Sharing information as desired, Perceiving a flowing and balanced interaction, Feeling positive emotions and Not having to engage coping mechanisms*. It also seems that participants consider *Being able to listen easily*, *Being spoken to in a helpful way*, and *Sharing information as desired* as significantly more important in group than in one-to-one conversations. These findings provided the foundation for the design of Study 2 (Chapter 3).

The subjective perception of conversation success was then evaluated during a face-to-face laboratory experiment (Chapter 3) involving groups of four people (quartets) with normal and impaired hearing. First, I showed that ratings of success given during a conversation strongly correlate with individuals' post-hoc overall assessment of conversation success. Also, I showed that the seven themes of conversation success mentioned earlier, converge to a single factor, through a factor analysis using survey questions aligned with the seven themes. Second, I demonstrated that self-reported perception of conversation success in four-way conversations is impacted by background noise, by hearing status and by use of hearing aids (Chapter 4). Participants held conversations with and without hearing aids and in three levels of background noise: low, medium, and high. When participants with impaired hearing wore their hearing aids, they seemed to rate conversation success significantly higher compared to the conversations in which they were unaided. At the same time, we found no evidence that aiding affects the normal hearing participants' perception of conversation success. In the unaided conditions, individuals with normal hearing perceive conversations as more successful than those with impaired hearing. These results show that hearing aids help people with hearing loss perceive conversation success in a similar way to normal hearing participants.

7.2.2 Behavioural perspectives on conversation success

What behaviours are associated with the perception of conversation success (Chapter 5)? Communication behaviours and perceptions were analysed at two levels: individual-level and group-level. Individual-level behaviours are

behaviours that have been derived from a single person such as the time a person spent talking. Group-level behaviours are measurements derived from all four participants simultaneously such as the time in which nobody is talking (silence time). Findings showed that both vocal activity and head movements are associated with the perception of conversation success. In terms of vocal activity: more talking was associated with an increase in individual conversation success but only for participants with impaired hearing; higher levels of imbalance in speech proportions among participants as well as prolonged periods of silence were related to a decrease in the groups' perception of conversation success. These results suggest that while having the opportunity to speak may be an indicator of conversation success for individuals with impaired hearing, it may not hold the same significance for those with normal hearing. While not speaking for normal hearing participants may have been a choice, impaired hearing participants may have simply been unable to do so due to hearing difficulties; consequently, when they do engage, they perceive conversations as more successful. At the group level, the findings indicate that balanced conversations, where no one person dominates or is isolated, are preferable for overall conversation success.

In terms of head movement: in unaided conditions, nodding was related to an increase in individual ratings of conversation success. In higher background noise conditions (medium and high) the perception of conversation success increased with the amount of time all three listeners spend simultaneously orienting towards the speaker. Nodding could suggest engagement which is a component of conversation success reflected in the theme *Being engaged and accepted*. Looking at the speaker may also signal engagement, especially in higher background noise conditions indicating individuals' efforts to attend to the speaker.

7.2.3 Third-party observers' perspectives on conversation success

How do third-party observers judge conversation success (Chapter 6)?

Here I demonstrate that 1) it is possible to know if a conversation is successful without being part of it, especially when observers have access to both auditory and visual cues rather than visual only; and 2) observers judge if a conversation

is successful using cues that fit under the seven original themes of conversation she success obtained in Study 1. However, results suggested that from the observers' perspective, on top of the seven pre-existing themes, a conversation can also be judged as successful when participants are seen to be *Engaging effort positively* or *Bonding over difficulty*. All but one of the original themes (*Not having to engage coping mechanisms*) were reflected in the observers' responses. Observer perspectives complete the conceptualisation of conversation success extending it to nine unique facets (Figure 7.2).

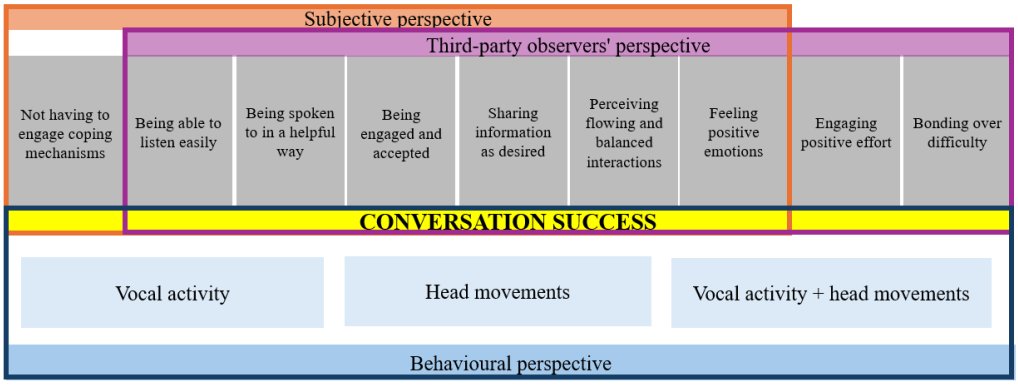


Figure 7.2 Summary of key findings on conversation success from three perspectives subjective (orange), behavioural (blue) and third-party observers (purple).

7.3 Theoretical and practical implications

These findings have implications across multiple domains. The conceptualisation of conversation success at a theoretical level could lay the groundwork for new and potentially enhanced theoretical models of human communication. At a practical level, these results could inform clinical practices and potentially enhance rehabilitation programs. Furthermore, for the hearing aid industry, these findings could inspire the development of innovative outcome measures for assessing the effectiveness of hearing aids. Ultimately, given that conversations are a fundamental aspect of daily life these findings have the potential to impact everyone interested in improving their interactions.

7.3.1 Theoretical models of human communication

The theoretical models of human communication described early on in this thesis (Chapter 1) do not take into account that human communication is initiated with the goal of achieving success. While conversation success is not always possible, it is an inherent aspiration. These theoretical models describe the structure of the conversation as if when all components of a conversation function optimally, the conversational process is successful. The findings presented in this thesis show that conversation success has multiple layers which can go beyond the components presented in existing theoretical models When mapping the conceptualisation of conversation success onto any described theoretical model, it becomes apparent that certain components align with established model elements. For instance, *Being able to listen easily* corresponds to the abilities of the "receiver" component, *Being spoken to in a helpful way* aligns with the characteristics of the "transmitter" component, and *Sharing information as desired* relates to the "message." However, other crucial aspects such as *Being engaged and accepted, Perceiving flowing and balanced interaction, Feeling positive emotions, Not having to engage coping mechanisms, Engaging positive effort*, and *Bonding over difficulty* are not reflected within any single component. The conceptualization of conversation success presented in this thesis has the potential to shape new theoretical models of human communication, which in turn could inspire innovative experimental paradigms.

Also, existent theoretical models of communication are limited to dyadic interaction. Through the second study (Chapters 3,4 and 5), this thesis shifts attention to group conversations. Moreover, Chapter 5 illustrates how the perception of group conversation success is related to group-level behaviours rather than solely individual behaviours. This highlights the need for theoretical models that expose the complexities of group communication dynamics.

7.3.2 Hearing aids

One of the aims of this thesis was to generate information that could potentially guide the development of hearing aids. Hearing aids are usually

tested in listening/comprehension-only tasks but are expected to provide benefits in conversational settings. However, listening-only tasks and engaging in conversation are distinct activities that require different testing approaches. One potential application of the findings of this thesis is to incorporate the themes of conversation success identified in Chapters 2 and 6 into hearing aid assessments, possibly through surveys tailored to social interaction. Furthermore, testing methods could be refined by incorporating behavioural indicators of conversation success identified in Study 2 (Chapter 5), such as head movements. With advancements in technology and machine learning, hearing aids could potentially detect head movements in real-time, allowing for adaptive adjustments that better accommodate group contexts and respond to individual listening needs. Although this thesis offers only an initial exploration of the behaviours linked to conversation success, it highlights promising avenues for future research dedicated to enhancing hearing aid technology in ways that support social interaction.

7.3.3 Clinical implications

Finally, the findings of this thesis might carry implications for clinical care in audiology. These results have the potential to assist audiologists in setting realistic expectations for individuals who are newly fitted with hearing aids. It might be common for people to anticipate that hearing aids will fully restore their ability to participate in conversations as before. Therefore, audiologists can use the conceptualisation of conversation success to educate patients about the potential outcomes of hearing aid use, helping them adjust their expectations accordingly. Similarly, the themes of conversation success could inspire aural rehabilitation programs to develop training for effective communication strategies. Additionally, involving frequent communication partners in these programs could enhance their communication skills and support individuals with hearing loss in reaching their communication goals.

7.4 General limitations and future directions

Like any research, the studies presented in this thesis have their limitations. One such limitation is related to the homogeneity of the participant

pool recruited for Studies 1 and 2. Participants were drawn from the Hearing Sciences Scottish Section participant pool and predominantly comprised Scottish, white Glasgow residents with access to medical care. Additionally, a significant portion of the participants were retired, potentially impacting the type, frequency, and environmental context of their interactions. This factor may introduce bias in comparison to individuals who are still actively employed. Furthermore, these participants agree to take part in experiments and are willing to do their best to help the research which could influence the way they act in the laboratory settings. Therefore, the generalizability of the findings may be constrained due to the specific demographic characteristics of the participant pool. Future research employing more diverse samples and longitudinal designs could offer insights into conversation success across varied populations and contexts.

Another limitation is represented by the exclusive focus on extra-linguistic behaviours. At times, extra-linguistic behaviour may be interpreted more accurately together with the verbal content, as demonstrated in Chapter 6. However, our behavioural analysis in Chapter 5 links the self-perception of conversation success only to extra-linguistic behaviours. While variability in conversation success perception could be explained by the verbal content, this question remains unanswered for future research. Similarly, another set of behaviours left to be investigated by future research pertains to facial expressions that could be extracted from the video recordings. Emotions, intentions, and even engagement could potentially be discerned through the assessment of the facial action units present during conversations (Cunningham et al., 2004; Nusseck et al., 2008). Study 2 (Chapter 3) has provided a wealth of data that holds potential for future research investigating behaviours not covered in this thesis. These could be extra-linguistic behaviours such as gestures (e.g.: hand waves), facial expressions, prosodic features (e.g.: intonation), or verbal content-related behaviour such as repetitions. Also, future research could go even further to additionally collect physiological data (e.g.: heart rate), brain activity (e.g.: hyper scanning) to relate to subjective experience and behaviour measures.

One area that could have been explored more thoroughly in this thesis is the contrast between individuals with normal hearing and those with impaired hearing. For example, in Study 1, the conceptualisation of conversation success could have been analysed separately for normal and impaired hearing participants, resulting in two distinct sets of themes. Similarly, in Study 2, conversation success could have been assessed in quartets involving only normal hearing participants or only impaired hearing participants to provide a clearer understanding of the differences. Additionally, in the third study, including impaired hearing observers could have allowed for a comparison between cues of success used by the two groups. Moving forward, future research could include this contrast to enhance our understanding of how conversation success varies based on hearing abilities.

Conversation behaviours can vary depending on factors such as the familiarity of participants (strangers/family), the nature of the interaction (social/transactional), and the context (at home/work). The studies presented in this thesis do not address these varied aspects but rather focus on conversation success overall. This omission leaves out a dimension that could be important in assessing conversation success. We are left wondering whether the requirements for a successful conversation differ based on the participants' relationship or the setting of the conversation. Future research is needed to address such questions and explore other dimensions of conversations.

The results presented in this thesis provide a solid foundation for future research. The nine themes of conversation success identified in Chapters 2 and 6 could inform the development of a standardised Conversation Success Questionnaire. Such a tool would be highly valuable for evaluating hearing aid performance in real-world social settings, complementing traditional listening-based assessments. Further research could also explore the factors influencing conversation success across various everyday contexts, such as the impact of conversational goals, group dynamics, or environmental noise levels on individuals' perceptions of conversation success.

This thesis also highlights specific behaviours linked to self-reported conversation success. For example, nodding was found to be associated with

more successful conversations, but its exact role remains unclear. Does nodding signal understanding and agreement while listening, or does it enhance clarity and engagement while speaking? Future research could examine combinations or sequences of behaviours, such as the differing impacts of nodding while speaking versus nodding while listening, to better understand how these behaviours collectively contribute to conversation success.

7.5 Conclusion

This thesis has focused on operationalising conversation success for both people with normal hearing and those with hearing loss. The ultimate motivation of my work is dedicated to making life easier for people with hearing loss, especially in conversations, which are vital for social connection and overall well-being. Because hearing loss can make it difficult to participate in conversations, it is essential to understand how hearing aids can help. The findings presented in this thesis shed light on the multifaceted nature of conversation success, demonstrating nine distinct themes and identifying potential behavioural indicators within multiparty interactions. While there is much left to explore, I hope the findings presented in this thesis spark further research in this area, leading to better ways of supporting people with hearing loss in their interactions.

Appendix

A: Study 1 Instructions

Brainstorming activity

Spoken conversation is the most common form of routine interaction. We engage in conversations every day: from chit-chatting with a stranger, discussions with work colleagues to meaningful conversations with our loved ones. Conversation is a complex activity. It can include many aspects, such as: talking, listening, understanding, gestures, eye-contact, facial expressions, and so on. Conversations can also be affected by the people involved and where they take place.

In this study we are interested in conversational success. The study will be conducted in three stages. In this first stage, you will be asked to write sentences related to conversational success. You can think about any type of face-to-face conversation you have had*. We encourage you to think about face-to-face conversations between two persons as well as larger groups of people. This survey takes approximately 30 minutes to complete. If you are currently busy or in a hurry, please come back to this page through the link in your email when it is more convenient for you to take part. Please feel free to take breaks during the survey if you need them, but please do not close this window before you are finished with it, as all your data will be lost.

*Currently, you may not be having as many face-to-face conversations as you had before the pandemic. If so, we encourage you to think back to the variety of conversational situations that are not happening now.

Please respond to the following questions. Please write a minimum of one statement, and a maximum of 10 statements for each question. Please use one box per statement. You can think about any type of face-to-face conversation you have had. We encourage you to think about face-to-face conversations between two persons as well as larger groups of people.

What does ''successful conversation'' look like?

***

Think about a successful conversation you have taken part in. What aspects of that conversation contributed to its success?

Sorting activity

Welcome to the 2nd Activity!

Successful conversations are hard to define. In this stage, you'll be given a list of cards that express potential factors of conversational success. You'll be asked to create groups of factors using the list. This task will take around 30 minutes.

Please enter below the token number received in the invitation email. Then, by clicking on "Begin" button, you will be directed to the instruction page.

***

On the left side of the page you will see a list with 73 cards. Each card represents a factor that can lead to conversational success.

The goal of this activity is to move the cards from the left and organize them into groups of similar items on the right. You will need to sort them in a way that makes sense to you. There is no right or wrong answer. Just do what comes naturally.

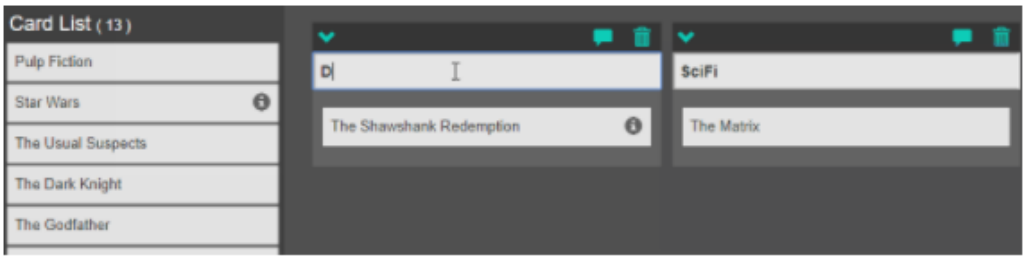
Please create at least 5 groups. A card could be put in its own group if it is unrelated to the other cards or if it stands alone as a unique idea. Please do not create a “Miscellaneous” or “Other” group.

Step 1: Read the sentences and drag* the cards from the left column to the right side of the page to create a new group.

**drag= click on the card, move the card on the right side by keeping the ‘’click’’ button pressed, release the button when you are above the area of the chosen group* .

Step 2: After you create a group you can rename it by clicking in the area above the card near the top of the group box ('Enter Group Name" bar).

Below you can see an example:



Step 3: Drag cards into the best suited groups on the right. Be sure to categorize all the cards. No card should be left aside.

Step 4: Once you have categorized all cards, use the I'm Done button at the top of the page to Submit and end the card sort.

Note: You may access these instructions at any time from the link in the top right corner of this page.

Rating activity

Welcome to the rating phase!

In this phase, you'll be given a list of statements related to conversational success. You'll be asked to rate how important each statement is for you in order for a conversation to be successful. This activity will take you around 30 minutes.

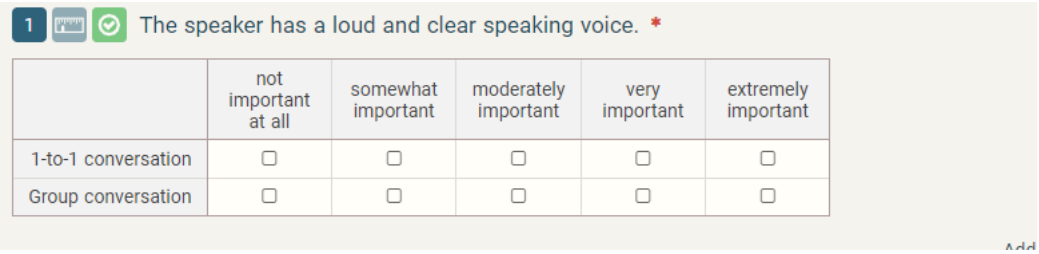
***

The number of people involved in a conversation can vary. You can have 1-to-1 conversations as well as conversations with larger groups of people.

On this page, you'll be asked to rate each statement thinking about their importance in conversations. Please tick a box in each row to assess how important each statement is for you, to make a conversation successful.

Certain statements may not be applicable to all participants and for those statements, there is an N/A option.

For you, how important are the following factors for a successful conversation?


1 The speaker has a loud and clear speaking voice. *

	not important at all	somewhat important	moderately important	very important	extremely important
1-to-1 conversation	☐	☐	☐	☐	☐
Group conversation	☐	☐	☐	☐	☐

B: Study 1 - List of statements

Number	Statement	Theme
1	The speaker has a loud and clear speaking voice.	Being spoken to in a helpful way
2	Being able to focus.	Being able to listen easily
3	Being able to follow gestures, lip movements and facial expressions.	Being able to listen easily
4	Participants clearly hearing what is being said.	Being able to listen easily
5	Body language reflecting engagement, e.g. nodding, smiling.	Being engaged and accepted
6	Passing and receiving information.	Sharing information as desired
7	A relaxed environment, with no pressure to "hurry" the conversation	Feeling positive emotions
8	Communicating a want or task.	Sharing information as desired
9	A balance between seriousness and humour.	Perceiving flowing and balanced interaction
10	Being able to fully understand the topic.	Being engaged and accepted
11	Sharing thoughts, feelings, experiences (past and/or future).	Feeling positive emotions
12	Participants understand the conversation.	Perceiving flowing and balanced interaction
13	Not having to strain to hear the other person(s).	Being able to listen easily
14	Not feeling anxious.	Not having to engage coping mechanisms
15	The speaker makes their points succinctly, without rambling.	Being spoken to in a helpful way
16	The speaker does not have to repeat themselves.	Being spoken to in a helpful way
17	The speaker doesn't get annoyed when I ask them to repeat themselves.	Being spoken to in a helpful way
18	The speaker is looking towards the listener(s) rather than around.	Being spoken to in a helpful way

19	The speaker is not making too many points in the one statement.	Being spoken to in a helpful way
20	The speaker's mouth isn't covered so the sound isn't obstructed (e.g.: mask/scarf /hand).	Being spoken to in a helpful way
21	Not needing to constantly adjust the volume on my hearing aid.	Being able to listen easily
22	Feeling part of the group.	Feeling positive emotions
23	Listener shows interest in what I'm talking about.	Being engaged and accepted
24	Hearing clearly, without having to assume what other the person was saying.	Being able to listen easily
25	Putting in minimal effort on my part.	Not having to engage coping mechanisms
26	Feeling positive emotions during the conversation (e.g.: relaxed, happy).	Feeling positive emotions
27	Leaving a conversation feeling pepped up.	Feeling positive emotions
28	Feeling joy in chatting with another human being.	Feeling positive emotions
29	Feeling the conversation was worth the time invested.	Feeling positive emotions
30	Feeling useful and appreciated (e.g.: helpful, valued).	Feeling positive emotions
31	Leaving a conversation feeling inspired.	Feeling positive emotions
32	Feeling comfortable when asking or being asked for something to be repeated.	Being engaged and accepted
33	Hearing aids working (e.g: not whistling, coming loose or batteries going flat).	Being able to listen easily
34	Feeling confident hearing due to functioning hearing aids.	Being able to listen easily
35	Having a to and fro debate process.	Perceiving flowing and balanced interaction
36	Exchanging feelings of empathy and compassion.	Feeling positive emotions
37	Working technology (when needed) that aids communication.	Being able to listen easily

38	The speaker talks fluently.	Being spoken to in a helpful way
39	Letting the topics flow naturally.	Perceiving flowing and balanced interaction
40	Not being distracted by other thoughts.	Sharing information as desired
41	Participants are laughing and being funny.	Feeling positive emotions
42	Participants are fully engaged and contributing.	Perceiving flowing and balanced interaction
43	Participants don't interrupt or talk over the top of each other.	Perceiving flowing and balanced interaction
44	Participants are polite, diplomatic and do not hurt feelings.	Perceiving flowing and balanced interaction
45	Feeling comfortable in the surroundings.	Not having to engage coping mechanisms
46	Feeling that you have learned something new at the end of the conversation.	Being engaged and accepted
47	No distractions in the background.	Being able to listen easily
48	The conversation flows easily.	Feeling positive emotions
49	Not finding myself withdrawing from the conversation.	Not having to engage coping mechanisms
50	A balance between asking questions and answering them.	Perceiving flowing and balanced interaction
51	Speaking with someone known to me as I can follow their voice better.	Not having to engage coping mechanisms
52	The group being small enough for me to be able to hear everyone.	Being able to listen easily
53	Other participants being aware of my hearing problem.	Not having to engage coping mechanisms
54	Being able to clearly communicate with professionals (e.g.: doctor).	Sharing information as desired
55	All participants contribute equally.	Perceiving flowing and balanced interaction
56	Regular eye contact is being made.	Not having to engage coping mechanisms
57	Having some kind of familiarity or friendship between the participants.	Perceiving flowing and balanced interaction

58	There is mutual knowledge about the topic of conversation.	Perceiving flowing and balanced interaction
59	All parties take the time to listen and to respond in turn.	Perceiving flowing and balanced interaction
60	The speaker talks at a steady pace rather than too quickly.	Being spoken to in a helpful way
61	Feeling professional.	Sharing information as desired
62	Being able to listen without trying to block out background noise.	Being able to listen easily
63	Not getting frustrated when the conversation is confusing.	Being engaged and accepted
64	Participants agree on an objective or an action.	Perceiving flowing and balanced interaction
65	Being in a quiet place without background noise or music.	Being able to listen easily
66	Positive encouragement is being exchanged.	Perceiving flowing and balanced interaction
67	Not having any pressure regarding the outcome(s) of the conversation.	Feeling positive emotions
68	It is a subject we all enjoy.	Perceiving flowing and balanced interaction
69	Personal opinions are listened to, even if conflicted.	Perceiving flowing and balanced interaction
70	Being in an environment with good acoustics.	Being able to listen easily
71	Sharing a sense of humor and/or irony.	Feeling positive emotions
72	Achieving a desired outcome.	Sharing information as desired
73	Having feedback during the conversation.	Sharing information as desired

C: Study 2 – Topics

Training conversation topic:

'Discuss how Glasgow has changed over the past ten years.'

Experimental conversation topics:

A. *'Discuss and plan a dinner party together using only food and drinks you all dislike.*
B. *'Imagine that you are writing a book together. Each of you is a character. Discuss what your character would be and what the book would be about.'*
C. *'Imagine that life on land is no longer possible. In order to survive you either have to live underwater or in space. Discuss which you would prefer as a group.'*
D. *'Imagine that you are heroes in a movie. What superpowers do you have and what will the movie be called?'*
E. *'Discuss and decide as a group what is the weirdest thing you have ever eaten. Would those in the group, who have never eaten that food, give it a try?'*
F. *'A genie has granted your group three wishes. Discuss and agree on which three wishes would be the best for all of you.'*
G. Extra topic used only in one quartet: '*Think about what movies you have watched in the past six months. Discuss and agree as a group on which are the best three movies.'*

D: Study 3 – Instructions

Instructions Prolific (the text that appears on participants' dashboard on prolific platform):

Conversation is a complex activity, and we want to know what makes it successful. Therefore, in this study, we are interested in **conversation success**.

To take part in this study you must:

- Be able to fluently read, write and comprehend English.
- Have self-perceived "normal" hearing and (corrected) vision.
- Have access to a computer or tablet with audio.

This survey will take approximately 30 minutes to complete.

You will watch 8 videos of people having conversations and answer questions about their interactions.

Wording in NOTE survey:

Spoken conversation is the most common form of routine interaction. We engage in conversations every day: from chit-chatting with a stranger and discussions with work colleagues to meaningful conversations with our loved ones. Conversation is a complex activity, and we want to know what makes it successful. Therefore, in this study, we are interested in **conversation success**.

You are going to watch 8 videos of group conversations. Each video is 2 minutes long. These videos were recorded in laboratory settings. That is why participants are equipped with unusual devices such as headsets and microphones. Try to ignore these elements and focus on their conversation. Please rate how successful you think the conversation is. Do this by using the rating scale at the bottom of the screen. There will be follow-up questions after each video.

Please listen to the following videos **in a quiet room without distractions**. It is recommended that you **use headphones**.

Please be aware that the audio will differ from video to video. Half of the videos will have no sound at all, and some may have conversations taking place in noisy background settings.

***

This is a practice video.

There will be audio so ensure your volume is at a level where you can hear the conversation easily.

Please watch the following clip. During the conversation, use the rating scale below the video when you perceive a change in conversation success. Move the pointer along the rating scale at least 3 times during the clip to rate how successful you think the conversation is.

***

Part 1

You will now start the first part of the study. You will be presented with 6 videos of group conversations to watch. Each video will take 2 minutes. Similar to the previous practice task, you will be asked to rate how successful you think the conversation is while you are watching each video. There will be some follow-up questions after each video.

Watch the following conversation. Please be aware there will be NO SOUND in this clip. Remember to use the rating scale when you perceive a change in conversation success. Once you have watched the clip, please answer the questions below with as much detail as possible.

***

How would you rate this conversation? *Response options from 1 to 5 (1= unsuccessful; 5 = successful)*

Successful moments: During the video, there may have been moments when you felt the conversation was progressing well. What specific features or actions led you to perceive these as successful instances in the conversation?

Unsuccessful moments*:* Conversely, there may have been moments when you felt the conversation was not going as well. What features or actions caused you to perceive these as less successful instances in the conversation?

***

Part 2

You will now start the second, and final, part of the study. You will watch 2 videos of group conversations. Each video will take 2 minutes. Similar to the previous practice task, you will be asked to rate how successful you think the conversation is while you are watching the video. There will be some follow-up questions after each video. Be aware that the follow-up questions in this part may differ from the questions you had previously.

Watch the following conversation. Please be aware there will be NO SOUND in this clip. Remember to use the rating scale when you perceive a change in

conversation success. Once you have watched the clip, please answer the questions below with as much detail as possible.

***

How would you rate this conversation? *Response options from 1 to 5 (1= unsuccessful; 5 = successful)*

Data from the questions below are not included in this thesis.

Some of the people you have just watched might have hearing loss. Based on your observation, please indicate which individuals (if any) you suspect to have hearing loss. *Response options (multiple options possible):*

- Person A
- Person B
- Person C
- Person D
- Nobody
- I don’t know

If selected any or more of the options Person A, B, C, D the following open question will appear:

Why? Please give as much detail as possible.

***

Thank you for your participation in this study.
You have come to the end of this survey. To register that you have completed the study, please copy the completion code below, return to Prolific.co. and click 'Finished Study', then paste this code into the box provided.

Completion code: XXXXXXXXX

E: Study 3 – List of words coded automatically

Frequent words and derivates	Successful AV	Unsuccessful AV	Successful V	Unsuccessful V
Engaged	X	X	X	X
Smile	X		X	X
Laughter	X		X	
Interested	X	X	X	X
Contribution	X	X	X	
Nods	X		X	
Join	X	X	X	X
Flow	X	X	X	
Animated	X		X	
Enjoyment	X		X	
Involvement	X	X	X	X
Gestures	X		X	
Happiness	X		X	
Relaxed	X			
Agreement			X	
Dominating		X		X
Bored		X		X
Shaking				X
Struggling		X		X
Leaning				X
Disengaged				X
Disinterested		X		X
Hearing (Heard)		X		
Difficult		X		
Pauses		X		
Awkward		X		
Silence		X		
Stilted		X		
Lulls		X		
Distracted		X		
Disagreement		X		

F: Study 3 – Top 30 most frequent words

Top 30 most frequent words used by observers for each condition and highlighted words that carry informational content about what observers were seeing or interpreting.

Successful moments				Unsuccessful moments			
V		AV		V		AV	
Word	**Count**	**Word**	**Count**	**Word**	**Count**	**Word**	**Count**
smiling	189	conversations	183	conversing	164	conversation	204
seemed	184	seemed	93	seemed	160	hearing	145
conversing	173	people	88	looking	151	one	131
looks	107	engaging	82	one	143	seemed	113
laughing	105	smiling	77	lady	80	people	71
engaging	99	laughing	71	man	55	talking	69
nods	93	everyone	55	moments	55	looking	64
lots	92	well	46	top	52	others	63
people	84	talking	43	left	50	noise	60
everyone	80	one	42	people	48	lady	57
laughter	68	group	41	talking	47	background	49
hand	66	interested	40	time	46	man	49
gesturing	58	laughter	40	appeared	45	speaking	47
animated	47	listening	37	others	45	like	42
talking	47	good	36	speaking	45	times	41
one	44	trying	34	hearing	44	bit	38
joining	42	contribution	33	dominating	40	group	38
listening	42	discussion	33	(not)engaged	39	person	36
agreement	41	asking	32	person	38	appeared	33
involvement	41	like	32	two	38	left	33
time	41	moments	31	unsuccessful	35	(not)engaged	32
appeared	38	nods	30	point	34	little	31
expressive	38	questions	30	like	33	much	31
interested	37	lots	29	much	33	moments	30
well	37	successful	28	none	30	struggling	29
lady	36	lady	28	(not) join	29	well	29
body	35	agreement	27	bit	27	asked	28
speaking	35	joining	27	bored	26	think	28
like	33	others	26	expressions	26	(no) interest	27

References

Abrams, H. B., Chisolm, T. H., & McArdle, R. (2005). Health-Related Quality of Life and Hearing Aids: A Tutorial. *Trends in Amplification*, *9*(3), 99–109. https://doi.org/10.1177/108471380500900302

Acheson, K. (2008). Silence as gesture: Rethinking the nature of communicative silences. *Communication Theory*, *18*(4), 535–555. https://doi.org/10.1111/j.1468-2885.2008.00333.x

Alarcón-Aldana, A. C., Callejas-Cuervo, M., & Bo, A. P. L. (2020). Upper limb physical rehabilitation using serious videogames and motion capture systems: A systematic review. *Sensors (Switzerland)*, *20*(21), 1–22. https://doi.org/10.3390/s20215989

Ambady, N., Bernieri, F., & Richeson, J. (2000). Toward a histology of social behavior: Thin-Slicing. *Advances in Experimental Social Psychology*, *32*, 201–271.

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology*, *83*(4), 947–961. https://doi.org/10.1037/0022-3514.83.4.947

Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., & Haynes, J. D. (2011). Flow of affective information between communicating brains. *NeuroImage*, *54*(1), 439–446. https://doi.org/10.1016/J.NEUROIMAGE.2010.07.004

Arlinger, S. (2003). Negative consequences of uncorrected hearing loss—a review. *International Journal of Audiology*, *42*(sup2), 17–20. https://doi.org/10.3109/14992020309074639

Aubanel, V., Cooke, M., Foster, E., Garcia Lecumberri, M. L., & Mayo, C. (2012). Effects of the availability of visual information and presence of competing conversations on speech production. *13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012*, *3*, 2031–2034. https://doi.org/10.21437/interspeech.2012-543

Aubanel, V., Cooke, M., Villegas, J., & Lecumberri, M. L. G. (2011).

Conversing in the presence of a competing conversation: Effects on speech production. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, August 2011*, 2833–2836. https://doi.org/10.21437/interspeech.2011-709

Bachorowski, J.-A., Smoski, M. J., & Owren, M. J. (2001). The acoustic features of human laughter. *The Journal of the Acoustical Society of America*, *110*(3), 1581–1597. https://doi.org/10.1121/1.1391244

Baker, R., & Hazan, V. (2011). DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, *43*(3), 761–770. https://doi.org/10.3758/s13428-011-0075-y

Bar, H., & Mentch, L. (2017). R-CMap—An open-source software for concept mapping. *Evaluation and Program Planning*, *60*. https://doi.org/10.1016/j.evalprogplan.2016.08.018

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians.[erratum appears in J Autism Dev Disord 2001 Dec;31(6):603]. *Journal of Autism & Developmental Disorders.*, *31*(1), 5–17. thesis

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beechey, T. (2023). Ordinal Pattern Analysis: A Tutorial on Assessing the Fit of Hypotheses to Individual Repeated Measures Data. *Journal of Speech, Language, and Hearing Research*, *66*(1), 347–364. https://doi.org/10.1044/2022_JSLHR-22-00133

Beechey, T., Buchholz, J. M., & Keidser, G. (2019). Eliciting naturalistic conversations: A method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment. *Journal of Speech, Language, and Hearing Research*, *62*(2), 470–484. https://doi.org/10.1044/2018_JSLHR-H-18-0107

Beechey, T., Buchholz, J. M., & Keidser, G. (2020). Hearing impairment

increases communication effort during conversations in noise. *Journal of Speech, Language, and Hearing Research*, *63*(1), 305–320. https://doi.org/10.1044/2019_JSLHR-19-00201

Bennett, Meyer, Eikelboom, & Atlas. (2018). Investigating the Knowledge, Skills, and Tasks Required for Hearing Aid Management: Perspectives of Clinicians and Hearing Aid Owners. *American Journal of Audiology*, *27*(1), 67–84. https://pubmed.ncbi.nlm.nih.gov/29305614/

Bennett, R. J., Saulsman, L., Eikelboom, R. H., & Olaithe, M. (2022). Coping with the social challenges and emotional distress associated with hearing loss: a qualitative investigation using Leventhal's self-regulation theory. *International Journal of Audiology*, *61*(5), 353–364. https://doi.org/10.1080/14992027.2021.1933620

Benus, S., Gravano, A., & Hirschberg, J. (2007). The Prosody of Backchannels in American English. *Proceedings of the 16th International Congress of Phonetic Sciences*, *August*, 1065–1068.

Berry, D. S., & Hansen, J. S. (1996). Positive Affect, Negative Affect, and Social Interaction. *Journal of Personality and Social Psychology*, *71*(4), 796–809. https://doi.org/10.1037/0022-3514.71.4.796

Beyan, C., Capozzi, F., Becchio, C., & Murino, V. (2018). Prediction of the leadership style of an emergent leader using audio and visual nonverbal features. *IEEE Transactions on Multimedia*, *20*(2), 441–456. https://doi.org/10.1109/TMM.2017.2740062

Bischoping, K. (1993). Gender differences in conversation topics, 1922-1990. *Sex Roles*, *28*(1–2), 1–18. https://doi.org/10.1007/BF00289744

Boholm, M., & Allwood, J. (2010). Repeated head movements, their function and relation to speech. *Proceedings of the Workshop on Multimodal Corpora Advances in Capturing Coding and Analyzing Multimodality LREC 2010*, 6–10. http://sskkii.gu.se/jens/publications/bfiles/B93.pdf

Bottalico, P., Piper, R. N., & Legner, B. (2022). Lombard effect, intelligibility, ambient noise, and willingness to spend time and money in a restaurant amongst older adults. *Scientific Reports*, *12*(1), 1–9. https://doi.org/10.1038/s41598-022-10414-6

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brimijoin, W. O., & Akeroyd, M. A. (2012). The role of head movements and signal spectrum in an auditory front/back illusion. *I-Perception*, *3*(3), 179–181. https://doi.org/10.1068/i7173sas

Brimijoin, W. O., McShefferty, D., & Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *The Journal of the Acoustical Society of America*, *127*(6), 3678–3688. https://doi.org/10.1121/1.3409488

Brooks, D. N., Hallam, R. S., & Mellor, P. A. (2001). The effects on significant others of providing a hearing aid to the hearing-impaired partner. *British Journal of Audiology*, *35*(3), 165–171. https://doi.org/10.1080/00305364.2001.11745234

Brumm, H., & Zollinger, A. (2011). The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour*, *148*(11–13), 1173–1198. https://doi.org/10.1163/000579511X605759

Burke, L. A., & Naylor, G. (2020). Daily-Life Fatigue in Mild to Moderate Hearing Impairment: An Ecological Momentary Assessment Study. *Ear and Hearing*, *41*(6), 1518–1532. https://doi.org/10.1097/AUD.0000000000000888

Burns, M. E., & Pearson, J. C. (2011). An exploration of family communication environment, everyday talk, and family satisfaction. *Communication Studies*, *62*(2), 171–185. https://doi.org/10.1080/10510974.2010.523507

Carlile, S., & Keidser, G. (2020). Conversational Interaction Is the Brain in Action: Implications for the Evaluation of Hearing and Hearing Interventions. *Ear and Hearing*, *41*, 56S-67S. https://doi.org/10.1097/AUD.0000000000000939

Chartrand, T. L., & Bargh, J. A. (1999). 1999_the_Chameleon_Effect.Pdf. In *Journal of Personality and Social Psychology* (Vol. 76, Issue 6, pp. 893–910).

Ciorba, A., Bianchini, C., Pelucchi, S., & Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical Interventions in Aging*, *7*, 159–163. https://doi.org/10.2147/CIA.S26059

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In *Advances in Psychology* (Vol. 9, Issue C). North-Holland Publishing Company. https://doi.org/10.1016/S0166-4115(09)60059-5

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process, Cognition, 22 (1986) l-39 1. *Cognition*, *22*, 1–39. https://pdfs.semanticscholar.org/dd2b/dd2c4df589cc3be1f4bfab6c42d8a9dc6609.pdf

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V., & Cowan, B. R. (2019). What makes a good conversation? Challenges in designing truly conversational agents. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3290605.3300705

Cooney, G., Mastroianni, A. M., Abi-Esber, N., & Brooks, A. W. (2020). The many minds problem: disclosure in dyadic versus group conversation. *Current Opinion in Psychology*, *31*, 22–27. https://doi.org/10.1016/j.copsyc.2019.06.032

Cunningham, D. W., Kleiner, M., Bülthoff, H. H., & Wallraven, C. (2004). The components of conversational facial expressions. *Proceedings - 1st Symposium on Applied Perception in Graphics and Visualization, APGV 2004*, *1*(212), 143–149. https://doi.org/10.1145/1012551.1012578

Cuperman, R., & Ickes, W. (2009). Big Five Predictors of Behavior and Perceptions in Initial Dyadic Interactions: Personality Similarity Helps Extraverts and Introverts, but Hurts "Disagreeables." *Journal of Personality and Social Psychology*, *97*(4), 667–684. https://doi.org/10.1037/a0015741

Dalton, D. S., Cruickshanks, K. J., Klein, B. E. K., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The Impact of Hearing Loss on Quality of Life in Older Adults. *Gerontologist*, *43*(5), 661–668.

https://doi.org/10.1093/geront/43.5.661

Danescu-Niculescu-mizil, C., Gamon, M., & Dumais, S. (2011). Mark my words! Linguistic style accommodation in social media. *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*, 745–754. https://doi.org/10.1145/1963405.1963509

Davis, M. H. (1980). Interpersonal Reactivity Index (IRI). *APA PsycTests*. https://doi.org/10.1037/t01093-000

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1–40. https://doi.org/10.1016/s0065-2601(01)80003-4

Drijvers, L., & Holler, J. (2022). Face-to-face spatial orientation fine-tunes the brain for neurocognitive processing in conversation. *IScience*, *25*(11), 105413. https://doi.org/10.1016/j.isci.2022.105413

Drimalla, H., Landwehr, N., Hess, U., & Dziobek, I. (2019). From face to face: the contribution of facial mimicry to cognitive and emotional empathy. *Cognition and Emotion*, *33*(8), 1672–1686. https://doi.org/10.1080/02699931.2019.1596068

Dunbar, R. I. M., Duncan, N. D. C., & Nettle, D. (1995). Size and structure of freely forming conversational groups. *Human Nature*, *6*(1), 67–78. https://doi.org/10.1007/BF02734136

Eagle, N., & Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, *10*(4), 255–268. https://doi.org/10.1007/s00779-005-0046-3

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). *Shared Reality Experiencing Commonality With Others' Inner States About the World*. *4*(5), 496–521.

Epley, N., & Schroeder, J. (2014). Mistakenly seeking solitude. *Journal of Experimental Psychology: General*, *143*(5), 1980–1999.

https://doi.org/10.1037/a0037323

Fay, N., Garrod, S., & Carletta, J. (2000). Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science*, *11*(6), 481–486. https://doi.org/10.1111/1467-9280.00292

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1477), 1669–1675. https://doi.org/10.1098/rspb.2001.1704

Floyd, K., Mikkelson, A. C., Tafoya, M. A., Farinelli, L., La Valley, A. G., Judd, J., Haynes, M. T., Davis, K. L., & Wilson, J. (2007). Human affection exchange: XIII. Affectionate communication accelerates neuroendocrine stress recovery. *Health Communication*, *22*(2), 123–132. https://doi.org/10.1080/10410230701454015

Foltz, A., Gaspers, J., Meyer, C., Thiele, K., Cimiano, P., & Stenneken, P. (2015). Temporal Effects of Alignment in Text-Based, Task-Oriented Discourse. *Discourse Processes*, *52*(8), 609–641. https://doi.org/10.1080/0163853X.2014.977696

Frick, R. W. (1985). Communicating Emotion : The Role of Prosodic Features. *Psychological Bulletin*, *97*(3), 412–429.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science*, *23*(8), 931–939. https://doi.org/10.1177/0956797612436816

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: A review. *Frontiers in Human Neuroscience*, *5*(FEBRUARY), 1–15. https://doi.org/10.3389/fnhum.2011.00011

Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes and Intergroup Relations*, *8*(2 SPEC. ISS.), 109–124.

https://doi.org/10.1177/1368430205051060

Garnier, M., Henrich, N., & Dubois, D. (2010). Influence of sound immersion and communicative interaction on the lombard effect. *Journal of Speech, Language, and Hearing Research*, *53*(3), 588–608. https://doi.org/10.1044/1092-4388(2009/08-0138)

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Gatica-Perez, D., McCowan, L., Zhang, D., & Bengio, S. (2005). Detecting group interest-level in meetings. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *I*(February 2005). https://doi.org/10.1109/ICASSP.2005.1415157

Glen, P. (2003). Laughter in Interaction. In *Cambridge: Cambridge University Press*.

Glista, D., O'Hagan, R., Moodie, S., & Scollie, S. (2020). An examination of clinical uptake factors for remote hearing aid support: a concept mapping study with audiologists. *International Journal of Audiology*, *60*(S1), S13–S22. https://doi.org/10.1080/14992027.2020.1795281

Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, *19*, 283–307.

Granberg, S., & Gustafsson, J. (2021). Key findings about hearing loss in the working-life: a scoping review from a well-being perspective. *International Journal of Audiology*, *60*(S2), 60–70. https://doi.org/10.1080/14992027.2021.1881628

Grebelsky-Lichtman, T. (2017). Verbal versus nonverbal primacy: Children's response to parental incongruent communication. *Journal of Social and Personal Relationships*, *34*(5), 636–661. https://doi.org/10.1177/0265407516651158

Grebelsky-Lichtman, T. (2021). Discrepant Verbal–Nonverbal Profile Theory. In *Engaging Theories in Interpersonal Communication* (3rd ed.). Routledge.

Grice, H. P. (1975). Logic and Conversation. *In Syntax and Semantics. Leiden, The Netherlands: Brill. Doi:*

Https://Doi.Org/10.1163/9789004368811_003.

Guydish, A. J., & Fox Tree, J. E. (2023). In Pursuit of a Good Conversation: How Contribution Balance, Common Ground, and Conversational Closings Influence Conversation Assessment and Conversational Memory. *Discourse Processes*, *60*(1), 18–41. https://doi.org/10.1080/0163853X.2022.2152552

Hadar, U., Steiner, T. J., Grant, E. C., & Clifford Rose, F. (1983). Head movement correlates of juncture and stress at sentence level. *Language and Speech*, *26*(2), 117–129. https://doi.org/10.1177/002383098302600202

Hadley, Fisher, & Pickering. (2020). Listeners are better at predicting speakers similar to themselves. *Acta Psychologica*, *208*. https://doi.org/10.1016/J.ACTPSY.2020.103094

Hadley, L. V., & Culling, J. F. (2022). Timing of head turns to upcoming talkers in triadic conversation: Evidence for prediction of turn ends and interruptions. *Frontiers in Psychology*, *13*(December), 1–14. https://doi.org/10.3389/fpsyg.2022.1061582

Hadley, L. V., Naylor, G., & Hamilton, A. F. de C. (2022). A review of theories and methods in the science of face-to-face social interaction. *Nature Reviews Psychology*, *1*(1), 42–54. https://doi.org/10.1038/S44159-021-00008-W

Hadley, L. V., Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2021a). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. *Psychonomic Bulletin and Review*, *28*(2), 632–640. https://doi.org/10.3758/s13423-020-01821-9

Hadley, L. V, Brimijoin, W. O., & Whitmer, W. M. (2019). Speech, movement, and gaze behaviours during dyadic conversation in noise. *Scientific Reports*, *9*(1), 1–8.

Hadley, L. V, Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2021b). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. *Psychonomic Bulletin & Review*, *28*(2), 632–640.

Hall, E. T. (1966). The Hidden Dimension. In *Anchor Books Editions* (Reprint. O, Issue 1).

Hall, J. A., & Davis, D. C. (2017). Proposing the Communicate Bond Belong Theory: Evolutionary Intersections With Episodic Interpersonal Communication. *Communication Theory*, *27*(1), 21–47. https://doi.org/10.1111/comt.12106

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159–177. https://doi.org/10.1007/s40167-016-0044-5

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences*, *18*, 83–90. https://doi.org/10.1016/j.cobeha.2017.07.018

Harmsen, I. E. (2019). Empathy in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(10), 3939–3955. https://doi.org/10.1007/s10803-019-04087-w

Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, *130*(4), 2139–2152. https://doi.org/10.1121/1.3623753

Healy, E. W., & Yoho, S. E. (2016). Difficulty understanding speech in noise by the hearing impaired: Underlying causes and technological solutions. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2016-Octob*, 89–92. https://doi.org/10.1109/EMBC.2016.7590647

Hecht, M. L. (1978). the Conceptualization and Measurement of Interpersonal Communication Satisfaction. *Human Communication Research*, *4*(3), 253–264. https://doi.org/10.1111/j.1468-2958.1978.tb00614.x

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. https://doi.org/10.1016/j.wocn.2010.08.002

Helvik, A. S., Krokstad, S., & Tambs, K. (2013). Hearing loss and risk of early

retirement. The HUNT study. *European Journal of Public Health*, *23*(4), 617–622. https://doi.org/10.1093/eurpub/cks118

Hepburn, A., Wilkinson, S., & Shaw, R. (2012). Repairing Self- and Recipient Reference. *Research on Language and Social Interaction*, *45*(2), 175–190. https://doi.org/10.1080/08351813.2012.673914

Hill, C. E., Thompson, B. J., & Ladany, N. (2003). Therapist use of silence in therapy: A survey. *Journal of Clinical Psychology*, *59*(4), 513–524. https://doi.org/10.1002/jclp.10155

Hirschi, Q., Wilson, T. D., & Gilbert, D. T. (2023). Speak Up! Mistaken Beliefs About How Much to Talk in Conversations. *Personality and Social Psychology Bulletin*, *49*(10), 1454–1465. https://doi.org/10.1177/01461672221104927

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma, D., & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-short). *Journal of Autism and Developmental Disorders*, *41*(5), 589–596. https://doi.org/10.1007/s10803-010-1073-0

Hogan, A., O'Loughlin, K., Davis, A., & Kendig, H. (2009). Hearing loss and paid employment: Australian population survey findings. *International Journal of Audiology*, *48*(3), 117–122. https://doi.org/10.1080/14992020802449008

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006

Holler, J., & Wilkin, K. (2011). *Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue*. https://doi.org/10.1007/s10919-011-0105-6

Holman, J. A., Ali, Y. H. K., & Naylor, G. (2023). A qualitative investigation of the hearing and hearing-aid related emotional states experienced by adults with hearing loss. *International Journal of Audiology*, *62*(10), 973–982. https://doi.org/10.1080/14992027.2022.2111373

Holman, J. A., Drummond, A., Hughes, S. E., & Naylor, G. (2019). Hearing impairment and daily-life fatigue: a qualitative study. *International Journal of Audiology*, *58*(7), 408–416. https://doi.org/10.1080/14992027.2019.1597284

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

Huang, K., Yeomans, M., Brooks, A. W., Minson, J., & Gino, F. (2017). It doesn't hurt to ask: Question-asking increases liking. *Journal of Personality and Social Psychology*, *113*(3), 430–452. https://doi.org/10.1037/pspi0000097

Hygge, S., Ronnberg, J., Larsby, B., & Arlinger, S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *Journal of Speech and Hearing Research*, *35*(1), 208–215. https://doi.org/10.1044/jshr.3501.208

IEEE Recommended Practice for Speech Quality Measurements. (1969). *IEEE Transactions on Audio and Electroacoustics*, *17*(3), 225–246. https://doi.org/10.1109/IEEESTD.1969.7405210

Intille, S., Haynes, C., Maniar, D., Ponnada, A., & Manjourides, J. (2016). μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1124–1135. https://doi.org/10.1145/2971648.2971717

Ishi, C. T., Ishiguro, H., & Hagita, N. (2014). Analysis of relationship between head motion events and speech in dialogue conversations. *Speech Communication*, *57*, 233–243. https://doi.org/10.1016/j.specom.2013.06.008

ISO. (2003). *Ergonomic assessment of speech communication*. *Geneva*(9921–1).

Jakobson, R. (1972). Motor signs for 'Yes' and 'No.' *Language in Society*, *1*(1), 91–96. https://doi.org/10.1017/S0047404500006564

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a “standard maximum” silence of approximately one second in conversation. In *Conversation: An interdisciplinary pespective* (pp. 166–196).

Jensen, M. (2016). Personality traits and nonverbal communication patterns. *International Journal of Social Science Studies*, *4*(5). https://doi.org/10.11114/ijsss.v4i5.1451

Joore, M. A., Potjewijd, J., Timmerman, A. A., & Anteunis, L. J. C. (2002). Response shift in the measurement of quality of life in hearing impaired adults after hearing aid fitting. *Quality of Life Research*, *11*(4), 299–307. https://doi.org/10.1023/A:1015598807510

Kagan, A., Winckel, J., Black, S., Duchan, J. F., Simmons-Mackie, N., & Square, P. (2004). A Set of Observational Measures for Rating Support and Participation in Conversation between Adults with Aphasia and Their Conversation Partners. *Topics in Stroke Rehabilitation*, *11*(1), 67–83. https://doi.org/10.1310/CL3V-A94A-DE5C-CVBE

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, *146*(4), 569–576. https://doi.org/10.1037/xge0000271

Keidser, G., Naylor, G., Brungart, D. S., Caduff, A., Campos, J., Carlile, S., Carpenter, M. G., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., & Smeds, K. (2020). The Quest for Ecological Validity in Hearing Science: What It Is, Why It Matters, and How to Advance It. *Ear and Hearing*, *41*, 5S-19S. https://doi.org/10.1097/AUD.0000000000000944

Kemper, S. (1994). Elderspeak: Speech Accommodations to Older Adults. *Aging, Neuropsychology, and Cognition*, *1*(1), 17–28. https://doi.org/10.1080/09289919408251447

Kendon, A. (2002). Some uses of the head shake. *Gesture*, *2*(2), 147–182. https://doi.org/10.1075/gest.2.2.03ken

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S.,

Chisolm, T., Davis, A. C., Erber, N. P., Hickson, L., Holmes, A., Rosenhall, U., & Von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *International Journal of Audiology*, *42*(SUPPL. 2). https://doi.org/10.3109/14992020309074650

Kohler, K. J. (2008). "Speech-smile", "speech-laugh", "laughter" and their sequencing in dialogic interaction. *Phonetica*, *65*(1–2), 1–18. https://doi.org/10.1159/000130013

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2011). Disrupting the flow: How brief silences in group conversations affect social needs. *Journal of Experimental Social Psychology*, *47*(2), 512–515. https://doi.org/10.1016/j.jesp.2010.12.006

Koutsombogera, M., & Vogel, C. (2022). Understanding Laughter in Dialog. *Cognitive Computation*, *14*(4), 1405–1420. https://doi.org/10.1007/s12559-022-10013-7

Krepsz, V., Horváth, V., Huszár, A., Neuberger, T., & Gyarmathy, D. (2024). 'Should we laugh?' Acoustic features of (in)voluntary laughters in spontaneous conversations. *Cognitive Processing*, *25*(1), 89–106. https://doi.org/10.1007/s10339-023-01168-8

Kumano, S., Ishii, R., & Otsuka, K. (2017). Comparing empathy perceived by interlocutors in multiparty conversation and external observers. *2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, *2018-Janua*, 50–57. https://doi.org/10.1109/ACII.2017.8273578

Lavan, N., Scott, S. K., & McGettigan, C. (2016). Laugh Like You Mean It: Authenticity Modulates Acoustic, Physiological and Perceptual Properties of Laughter. *Journal of Nonverbal Behavior*, *40*(2), 133–149. https://doi.org/10.1007/s10919-015-0222-8

Lee, C. C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., Georgiou, P. G., & Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. *Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010*, *January*,

793–796. https://doi.org/10.21437/interspeech.2010-287

Leertouwer, I., Schuurman, N. K., & Vermunt, J. K. (2022). Are Retrospective Assessments Means of People's Experiences? Accounting for Interpersonal and Intrapersonal Variability when Comparing Retrospective Assessment Data to Ecological Momentary Assessment Data. *Journal for Person-Oriented Research*, *8*(2), 52–70. https://doi.org/10.17505/jpor.2022.24855

Lelic, D., Wolters, F., & Schinkel-Bielefeld, N. (2023). Journal of the American Academy of Audiology Measuring hearing aid satisfaction in everyday listening situations: Retrospective and in-situ assessments complement each other. *Journal of the American Academy of Audiology*. https://doi.org/10.1055/a-2265-9418

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). *Package "Emmeans" R Package Version 4.0-3*. http://cran.r-project.org/package=emmeans

Levinson, S. C. (2020). On the human "interaction engine." *Roots of Human Sociality: Culture, Cognition and Interaction*, 39–69. https://doi.org/10.4324/9781003135517-3

Levinson, S. C., Holler, J., & Levinson, S. C. (2014). *The origin of human multi-modal communication*.

Levitan, R., Gravano, A., Willson, L., Beňuš, Š., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. *NAACL HLT 2012 - 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, 11–19.

Li, C. M., Zhang, X., Hoffman, H. J., Cotch, M. F., Themann, C. L., & Wilson, M. R. (2014). Hearing impairment associated with depression in US adults, National Health and Nutrition Examination Survey 2005-2010. *JAMA Otolaryngology - Head and Neck Surgery*, *140*(4), 293–302. https://doi.org/10.1001/jamaoto.2014.42

Lind, C. (2013). Conversation repair: Ecological validity of outcome measures in acquired hearing impairment. *Cochlear Implants International*,

14(SUPPL. 4). https://doi.org/10.1179/1467010013Z.000000000135

Lind, C., Hickson, L., & Erber, N. (2010). Who said what? Sampling conversation repair behavior involving adults with acquired hearing impairment. *Seminars in Hearing*, *31*(2), 104–115. https://doi.org/10.1055/s-0030-1252104

Lind, C., Hickson, L., & Erber, N. P. (2004). Conversation repair and acquired hearing impairment: A preliminary quantitative clinical study. *Australian and New Zealand Journal of Audiology*, *26*(1), 40–52. https://doi.org/10.1375/audi.26.1.40.55987

Lind, C., Hickson, L., & Erber, N. P. (2006). Conversation repair and adult cochlear implantation: A qualitative case study. *Cochlear Implants International*, *7*(1), 33–48. https://doi.org/10.1002/cii.20

Littlejohn and Foss. (2011). *Theories of human communication - communication theory and scholarship.pdf*.

Liu, M., Duan, Y., Ince, R. A. A., Chen, C., Garrod, O. G. B., Schyns, P. G., & Jack, R. E. (2022). Facial expressions elicit multiplexed perceptions of emotion categories and dimensions. *Current Biology*, *32*(1), 200-209.e6. https://doi.org/10.1016/j.cub.2021.10.035

Louven, C., Scholle, C., Gehrs, F., & Lenz, A. (2022). emoTouch Web – a web-based system for continuous real time studies with smartphones, tablets, and desktop computers. *Jahrbuch Musikpsychologie*, *30*(April). https://doi.org/10.5964/jbdgm.137

Manley, G. A. (2017). Comparative Auditory Neuroscience: Understanding the Evolution and Function of Ears. *JARO - Journal of the Association for Research in Otolaryngology*, *18*(1), 1–24. https://doi.org/10.1007/s10162-016-0579-3

Manning, J. (2014). A Constitutive Approach to Interpersonal Communication Studies. *Communication Studies*, *65*(4), 432–440. https://doi.org/10.1080/10510974.2014.927294

Maynard, S. K. (1987). Interactional functions of a nonverbal sign Head movement in japanese dyadic casual conversation. *Journal of Pragmatics*, *11*(5), 589–606. https://doi.org/10.1016/0378-2166(87)90181-0

McCaffrey, S. A., Chiauzzi, E., Chan, C., & Hoole, M. (2018). Understanding 'Good Health care' from the Patient's Perspective: Development of a Conceptual Model Using Group Concept Mapping. *The Patient - Patient-Centered Outcomes Research 2018 12:1*, *12*(1), 83–95. https://doi.org/10.1007/S40271-018-0320-X

McFarland, D. A., Jurafsky, D., & Rawlings, C. (2013). Making the connection: Social bonding in courtship situations. *American Journal of Sociology*, *118*(6), 1596–1649. https://doi.org/10.1086/670240

McLaughlin, M. L., & Cody, M. J. (1982). Awkward Silences: Behavioral Antecedents and Consequences of the Conversational Lapse. *Human Communication Research*, *8*(4), 299–316. https://doi.org/10.1111/j.1468-2958.1982.tb00669.x

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, *49*(2), 201–213. https://doi.org/10.1016/S0749-596X(03)00028-7

Meyer, A. S. (2023). Timing in Conversation. *Journal of Cognition*, *6*(1), 1–17. https://doi.org/10.5334/joc.268

Meyer, C., Waite, M., Atkins, J., Ekberg, K., Scarinci, N., Barr, C., Cowan, R., & Hickson, L. (2021). How Can eHealth Meet the Hearing and Communication Needs of Adults With Hearing Impairment and their Significant Others? A Group Concept Mapping Study. *Ear & Hearing*, *Publish Ah*, 1–12. https://doi.org/10.1097/aud.0000000000001097

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, *20*(3), 336–366. https://doi.org/10.1111/josl.1_12177

Monnet, T., Samson, M., Bernard, A., David, L., & Lacouture, P. (2014). Measurement of three-dimensional hand kinematics during swimming with a motion capture system: A feasibility study. *Sports Engineering*, *17*(3), 171–181. https://doi.org/10.1007/s12283-014-0152-4

Moore, H. T. (1992). Further data concerning sex differences. *The Journal of Abnormal Psychology and Social Psychology*, *17(2)*, 210–214.

https://doi.org/10.1037/h0064645

Morgan, A., Hickson, L., & Worrall, L. (2002). The impact of hearing impairment on quality of life of older people. *Asia Pacific Journal of Speech, Language and Hearing*, *7*(1), 39–53. https://doi.org/10.1179/136132802805576562

Most, T., & Aviner, C. (2009). Auditory, visual, and auditory - Visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, *14*(4), 449–464. https://doi.org/10.1093/deafed/enp007

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory-visual perception of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research*, *55*(1148–1162).

Most, T., Weisel, A., & Zaychik, A. (1993). Auditory, visual and auditory—visual identification of emotions by hearing and hearing-impaired adolescents. *British Journal of Audiology*, *27*(4), 247–253.

Mulwa, K. W., & Kucker, S. C. (2024). Coding social interactions in naturalistic settings: The taxonomy of dyadic conversation. *Behavior Research Methods*, *56*(1), 172–186. https://doi.org/10.3758/s13428-022-02033-w

Murphy, N. A., & Hall, J. A. (2021). Capturing Behavior in Small Doses: A Review of Comparative Research in Evaluating Thin Slices for Behavioral Measurement. *Frontiers in Psychology*, *12*(April), 1–13. https://doi.org/10.3389/fpsyg.2021.667326

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 169–172. https://doi.org/10.3115/1557690.1557737

Niederehe, G., & Duncan, S. (1974). On Signalling That it's Your Turn to Speak. *Journal of Experimental Social Psychology*, *10*(3), 234―247.

Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable

for clinical use: The SSQ12. *International Journal of Audiology*, *52*(6), 409–412. https://doi.org/10.3109/14992027.2013.781278.A

Nölle, J., Chen, C., Hensel, L. B., Garrod, O. G. B., Schyns, P. G., & Jack, R. E. (2021). Facial expressions of emotion include iconic signals of rejection and acceptance. *Journal of Vision*, *21*(9), 2932. https://doi.org/10.1167/jov.21.9.2932

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision*, *8*(8), 1–23. https://doi.org/10.1167/8.8.1

Oertel, C., Scherer, S., & Campbell, N. (2011). On the use of multimodal cues for the prediction of degrees of involvement in spontaneous conversation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, January 2014*, 1541–1544. https://doi.org/10.21437/interspeech.2011-465

Okada, S., Ohtake, Y., Nakano, Y. I., Hayashi, Y., Huang, H. H., Takase, Y., & Nitta, K. (2016). Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets. *ICMI 2016 - Proceedings of the 18th ACM International Conference on Multimodal Interaction*, *1*, 169–176. https://doi.org/10.1145/2993148.2993154

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2011). Best Practices in Exploratory Factor Analysis. In *Best Practices in Quantitative Methods* (Issue October). https://doi.org/10.4135/9781412995627.d8

Ostrand, R., & Chodroff, E. (2021). It's alignment all the way down, but not all the way up: Speakers align on some features but not others within a dialogue. *Journal of Phonetics*, *88*, 101074. https://doi.org/10.1016/j.wocn.2021.101074

Otsuka, K., & Tsumori, M. (2020). Analyzing Multifunctionality of Head Movements in Face-to-Face Conversations Using Deep Convolutional Neural Networks. *IEEE Access*, *8*, 217169–217195. https://doi.org/10.1109/ACCESS.2020.3041672

Palmer, A. D., Carder, P. C., White, D. L., Saunders, G., Woo, H., Graville, D.

J., & Newsom, J. T. (2019). The impact of communication impairments on the social relationships of older adults: Pathways to psychological well-being. *Journal of Speech, Language, and Hearing Research*, *62*(1), 1–21. https://doi.org/10.1044/2018_JSLHR-S-17-0495

Paluch, R., Latzel, M., Meis, M., & RICHARD PALUCH MATTHIAS LATZEL AND MARKUS MEIS. (2015). A new tool for subjective assessment of hearing aid performance: Analyses of Interpersonal Communication. *Proceedings of the International Symposium on Auditory and Audiological Research*, *5*(August), 453–460.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, *119*(4), 2382–2393. https://doi.org/10.1121/1.2178720

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2018). Similar neural responses predict friendship. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02722-7

Paxton, A., & Dale, R. (2017). Interpersonal movement synchrony responds to high- and low-level conversational constraints. *Frontiers in Psychology*, *8*(JUL), 1–16. https://doi.org/10.3389/fpsyg.2017.01135

Peperkoorn, L. S., Becker, D. V., Balliet, D., Columbus, S., Molho, C., & Van Lange, P. A. M. (2020). The prevalence of dyads in social life. *PLoS ONE*, *15*(12 December), 1–17. https://doi.org/10.1371/journal.pone.0244188

Petersen, E. B., MacDonald, E. N., & Josefine Munch Sørensen, A. (2022). The Effects of Hearing-Aid Amplification and Noise on Conversational Dynamics Between Normal-Hearing and Hearing-Impaired Talkers. *Trends in Hearing*, *26*. https://doi.org/10.1177/23312165221103340

Petersen, E., Walravens, El., & Kofoed Pedersen, A. (2023). Backchannel Behavior in Conversations and How it is Affected by Hearing Loss, Noise, and Hearing Aids. *10th Convention of the European Acoustics Association, September*, 2417–2424. https://doi.org/10.61782/fa.2023.0432

Picou, E. M. (2022). Hearing Aid Benefit and Satisfaction Results from the

MarkeTrak 2022 Survey: Importance of Features and Hearing Care Professionals. *Seminars in Hearing*, *43*(4), 301–316. https://doi.org/10.1055/s-0042-1758375

Poost-Foroosh, L., Jennings, M. B., Shaw, L., Meston, C. N., & Cheesman, M. F. (2011). Factors in Client–Clinician Interaction That Influence Hearing Aid Adoption. *Trends in Amplification*, *15*(3), 127–139. https://doi.org/10.1177/1084713811430217

Provine, R. . (1996). Contagious Yawning and Laughter: Significance for Sensory Feature Detection, Motor Pattern Generation, Imitation, and the Evolution of Social Behavior. In *Social Learning in Animals*. ACADEMIC PRESS, INC. https://doi.org/10.1016/b978-012273965-1/50010-8

Provine, R. R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, *30*(1), 1–4. https://doi.org/10.3758/BF03330380

Provine, R. R., & Fischer, K. R. (1989). Laughing, Smiling, and Talking: Relation to Sleeping and Social Context in Humans. *Ethology*, *83*(4), 295–305. https://doi.org/10.1111/j.1439-0310.1989.tb00536.x

Raman, C., Prabhu, N. R., & Hung, H. (2023). Perceived Conversation Quality in Spontaneous Interactions. *IEEE Transactions on Affective Computing*, *14*(4), 2901–2912. https://doi.org/10.1109/TAFFC.2023.3233950

Ramsberger, G., & Rende, B. (2002). Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, *16*(3), 337–353. https://doi.org/10.1080/02687040143000636

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2018). Daily well-being: The role of autonomy, competence, and relatedness. *Relationships, Well-Being and Behaviour: Selected Works of Harry Reis*, 317–349. https://doi.org/10.4324/9780203732496-13

Reitter, D., Keller, F., & Moore, J. D. (2011). A Computational Cognitive Model of Syntactic Priming. *Cognitive Science*, *35*(4), 587–637. https://doi.org/10.1111/j.1551-6709.2010.01165.x

Richmond, V. P., McCroskey, J. C., McCroskey, L. L., & Fayer, J. M. (2008).

Communication Traits In First And Second Languages. *Journal of Intercultural Communication Research*, *37*(2), 65–73. https://doi.org/10.1080/17475750802533331

Riest, C., Jorschick, A. B., & de Ruiter, J. P. (2015). Anticipation in turn-taking: Mechanisms and information sources. *Frontiers in Psychology*, *6*(FEB), 1–14. https://doi.org/10.3389/fpsyg.2015.00089

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247–261. https://doi.org/10.1080/14992027.2018.1551631

Rosas, S., & Kane, M. (2012). Quality and rigor of the concept mapping methodology: a pooled study analysis. *Evaluation and Program Planning*, *35*(2), 236–245. https://doi.org/10.1016/J.EVALPROGPLAN.2011.10.003

Rubin, R. B., Perse, E. M., & Barbato, C. A. (1988). Conceptualization and Measurement of Interpersonal Communication Motives I shape and nurture their self-concepts, make decisions about. *Human Communication Research International Communication Association*, *14*(4), 602–628.

Russell, D., Peplau, L., & Ferguson, M. L. (1978). UCLA Loneliness Scale. *APA PsycTests*. https://doi.org/10.1037/t05974-000

Sacks, H, Schegloff, E., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation Harvey Sacks , Emanuel A . Schegloff and Gail Jefferson. *Language*, *50*(4), 696–735.

Satterthwaite, F. E. (1946). An Approximate Distribution on Estimates of Variance Components. *Biometrics Bulletin*, *2*, 110–114. http://dx.doi.org/10.2307/3002019

Sauter, D. (2010). More than happy: The need for disentangling positive emotions. *Current Directions in Psychological Science*, *19*(1), 36–40. https://doi.org/10.1177/0963721409359290

Scarinci, N., Worrall, L., & Hickson, L. (2008). The effect of hearing

impairment in older people on the spouse. *International Journal of Audiology*, *47*(3), 141–151. https://doi.org/10.1080/14992020701689696

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. *Roundtable on Languages and Linguistics*, *March*, 71–93.

Schinkel-bielefeld, N. (2020). *Assessment : What Situations*. *29*(September), 591–609.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233. https://doi.org/10.1016/j.cognition.2009.08.009

See, A., Roller, S., Kiela, D., & Weston, J. (2019). What makes a good conversation? How controllable attributes affect human judgments. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 1702–1723. https://arxiv.org/abs/1902.08654v2

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shiell, M. M., Zaar, J., Høy-Christensen, J., Skoglund, M. A., Keidser, G., & Rotger-Griful, S. (2023). Multilevel Modeling of Gaze From Listeners With Hearing Loss Following a Realistic Conversation. *Journal of Speech, Language, and Hearing Research*, *66*(11), 4575–4589. https://doi.org/10.1044/2023_JSLHR-22-00641

Smeds, K., Gotowiec, S., Wolters, F., Herrlin, P., Larsson, J., & Dahlquist, M. (2020). *Smeds - Selecting_Scenarios_for_Hearing_Related_Laboratory.3-2*. 20–30.

Sørensen, A. J. M., Fereczkowski, M., & MacDonald, E. N. (2021). Effects of Noise and Second Language on Conversational Dynamics in Task Dialogue. *Trends in Hearing*, *25*. https://doi.org/10.1177/23312165211024482

Spitzberg, B. H., & Canary, D. J. (1985). *Loneliness and relationally*

competent communication (pp. 387–402). Journal of Social and Personal relationships, Vol. 2.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, *25*(5–6), 701–721. https://doi.org/10.1023/a:1020867916902

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society*, *35*(3), 367–392. https://doi.org/10.1017/S0047404506060179

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., & Kaplan, G. A. (2000). Negative Consequences of Hearing Impairment in Old AgeA Longitudinal Analysis. *The Gerontologist*, *40*(3), 320–326. https://doi.org/10.1093/GERONT/40.3.320

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *Journal of the Acoustical Society of America*, *26*(2), 212–215. https://doi.org/10.1121/1.1907309

Tanaka, H., & Campbell, N. (2011). Acoustic Features of Four Types of Laughter in Natural Conversational Speech. *International Congress of Phonetic Sciences*, *August*. https://api.semanticscholar.org/CorpusID:22893960

Templeton, E. M., Chang, L. J., Reynolds, E. A., Cone Lebeaumont, M. D., & Wheatley, T. (2023). Long gaps between turns are awkward for strangers but not for friends. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1875). https://doi.org/10.1098/rstb.2021.0471

ter Bekke, M., Drijvers, L., & Holler, J. (2024). Hand Gestures Have Predictive Potential During Conversation: An Investigation of the Timing of Gestures in Relation to Speech. *Cognitive Science*, *48*(1). https://doi.org/10.1111/cogs.13407

To, W. M., & Chung, A. W. L. (2015). Restaurant noise: Levels and temporal characteristics. *Noise and Vibration Worldwide*, *46*(8), 11–17. https://doi.org/10.1260/0957-4565.46.8.11

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12*(1), 1–16. https://doi.org/10.1016/0149-7189(89)90016-5

Trochim, W., Cook, J. A., & Setze, R. J. (1994). Using Concept Mapping to Develop a Conceptual Framework of Staff's Views of a Supported Employment Program for Individuals With Severe Mental Illness. *Journal of Consulting and Clinical Psychology*, *62*(4), 766–775. https://doi.org/10.1037/0022-006X.62.4.766

Trujillo, J., Özyürek, A., Holler, J., & Drijvers, L. (2021). Speakers exhibit a multimodal Lombard effect in noise. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-95791-0

Trujillo, J. P., & Holler, J. (2024). Conversational facial signals combine into compositional meanings that change the interpretation of speaker intentions. *Scientific Reports*, *14*(1), 1–14. https://doi.org/10.1038/s41598-024-52589-0

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*(NOV). https://doi.org/10.3389/FPSYG.2014.01323/ABSTRACT

Tye-Murray, N. (2003). Conversational fluency of children who use cochlear implants. *Ear and Hearing*, *24*(1 SUPPL.), 82–89. https://doi.org/10.1097/01.aud.0000051691.33869.ec

Vas, V., Akeroyd, M. A., & Hall, D. A. (2017). *A Data-Driven Synthesis of Research Evidence for Domains of Hearing Loss , as Reported by Adults With Hearing Loss and Their Communication Partners*. *21*, 1–25. https://doi.org/10.1177/2331216517734088

Ventry, I. M., & Weinstein, B. E. (1982). The hearing handicap inventory for the elderly: A new tool. *Ear and Hearing*, *3*(3), 128–134. https://doi.org/10.1097/00003446-198205000-00006

Vettin, J., & Todt, D. (2004). Laughter in conversation: Features of occurrence and acoustic structure. *Journal of Nonverbal Behavior*, *28*(2), 93–115. https://doi.org/10.1023/B:JONB.0000023654.73558.72

Vivek, Patil H., Singh, S. N., Mishra, S., & Donavan, D. T. (2007). *Parallel Analysis Engine to Aid Determining Number of Factors to Retain*. https://analytics.gonzaga.edu/parallelengine/

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Dilegge, T., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Fleming, T., Forouzanfar, M. H., … Zuhlke, L. J. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, *57*, 209–232. https://doi.org/10.1016/j.specom.2013.09.008

Warner-Garcia, S. (2014). Laughing when nothing's funny: The pragmatic use of coping laughter in the negotiation of conversational disagreement. *Pragmatics*, *24*(1), 157–180. https://doi.org/10.1075/prag.24.1.07war

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weick, K. E. (1979). *The Social Psychology of Organizing* (Second). McGraw-Hill (1979).

Whitmer, W. M., McShefferty, D., & Smeds, K. (2024). Exploring attentive listening in noise through the just-follow conversation task. *SPIN2024*, P04.

Wilson, J., Hickson, L., & Worrall, L. (1998). Use of communication strategies by adults with hearing impairment. *Asia Pacific Journal of Speech, Language and Hearing*, *3*(1), 29–42. https://doi.org/10.1179/136132898805577250

Wolters, F., Smeds, K., Schmidt, E., Christensen, E. K., & Norup, C. (2016). Common sound scenarios: A context-driven categorization of everyday sound environments for application in hearing-device research. *Journal of the American Academy of Audiology*, *27*(7), 527–540. https://doi.org/10.3766/jaaa.15105

World Health Organization. (2001). *ICF: International Classification of Functioning, Disability and Health*.

Wu, Y. H., Stangl, E., Chipara, O., Gudjonsdottir, A., Oleson, J., & Bentler, R. (2020). Comparison of In-Situ and Retrospective Self-Reports on Assessing Hearing Aid Outcomes. *Journal of the American Academy of Audiology*, *31*(10), 746–757. https://doi.org/10.1055/s-0040-1719133

Wynn, C. J., Barrett, T. S., & Borrie, S. A. (2022). Rhythm Perception, Speaking Rate Entrainment, and Conversational Quality: A Mediated Model. *Journal of Speech, Language, and Hearing Research*, *65*(6), 2187–2203. https://doi.org/10.1044/2022_JSLHR-21-00293

Yeomans, M., Schweitzer, M. E., & Brooks, A. W. (2022). The Conversational Circumplex: Identifying, prioritizing, and pursuing informational and relational motives in conversation. *Current Opinion in Psychology*, *44*, 293–302. https://doi.org/10.1016/j.copsyc.2021.10.001

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex*, *27*(10), 4988–5000. https://doi.org/10.1093/cercor/bhx202

Zevon, M. A., Schwabish, S., Donnelly, J. P., & Kerry, R. (2007). Medically related legal needs and quality of life in cancer care: a structural analysis. *Cancer*, *109*(12), 2600–2606. https://doi.org/10.1002/CNCR.22682

Abrams, H. B., Chisolm, T. H., & McArdle, R. (2005). Health-Related Quality of Life and Hearing Aids: A Tutorial. *Trends in Amplification*, *9*(3), 99–

109. https://doi.org/10.1177/108471380500900302

Acheson, K. (2008). Silence as gesture: Rethinking the nature of communicative silences. *Communication Theory*, *18*(4), 535–555. https://doi.org/10.1111/j.1468-2885.2008.00333.x

Alarcón-Aldana, A. C., Callejas-Cuervo, M., & Bo, A. P. L. (2020). Upper limb physical rehabilitation using serious videogames and motion capture systems: A systematic review. *Sensors (Switzerland)*, *20*(21), 1–22. https://doi.org/10.3390/s20215989

Ambady, N., Bernieri, F., & Richeson, J. (2000). Toward a histology of social behavior: Thin-Slicing. *Advances in Experimental Social Psychology*, *32*, 201–271.

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology*, *83*(4), 947–961. https://doi.org/10.1037/0022-3514.83.4.947

Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., & Haynes, J. D. (2011). Flow of affective information between communicating brains. *NeuroImage*, *54*(1), 439–446. https://doi.org/10.1016/J.NEUROIMAGE.2010.07.004

Arlinger, S. (2003). Negative consequences of uncorrected hearing loss—a review. *International Journal of Audiology*, *42*(sup2), 17–20. https://doi.org/10.3109/14992020309074639

Aubanel, V., Cooke, M., Foster, E., Garcia Lecumberri, M. L., & Mayo, C. (2012). Effects of the availability of visual information and presence of competing conversations on speech production. *13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012*, *3*, 2031–2034. https://doi.org/10.21437/interspeech.2012-543

Aubanel, V., Cooke, M., Villegas, J., & Lecumberri, M. L. G. (2011). Conversing in the presence of a competing conversation: Effects on speech production. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, August 2011*, 2833–2836. https://doi.org/10.21437/interspeech.2011-709

Bachorowski, J.-A., Smoski, M. J., & Owren, M. J. (2001). The acoustic features of human laughter. *The Journal of the Acoustical Society of America*, *110*(3), 1581–1597. https://doi.org/10.1121/1.1391244

Baker, R., & Hazan, V. (2011). DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, *43*(3), 761–770. https://doi.org/10.3758/s13428-011-0075-y

Bar, H., & Mentch, L. (2017). R-CMap—An open-source software for concept mapping. *Evaluation and Program Planning*, *60*. https://doi.org/10.1016/j.evalprogplan.2016.08.018

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians.[erratum appears in J Autism Dev Disord 2001 Dec;31(6):603]. *Journal of Autism & Developmental Disorders.*, *31*(1), 5–17. thesis

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beechey, T. (2023). Ordinal Pattern Analysis: A Tutorial on Assessing the Fit of Hypotheses to Individual Repeated Measures Data. *Journal of Speech, Language, and Hearing Research*, *66*(1), 347–364. https://doi.org/10.1044/2022_JSLHR-22-00133

Beechey, T., Buchholz, J. M., & Keidser, G. (2019). Eliciting naturalistic conversations: A method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment. *Journal of Speech, Language, and Hearing Research*, *62*(2), 470–484. https://doi.org/10.1044/2018_JSLHR-H-18-0107

Beechey, T., Buchholz, J. M., & Keidser, G. (2020). Hearing impairment increases communication effort during conversations in noise. *Journal of Speech, Language, and Hearing Research*, *63*(1), 305–320. https://doi.org/10.1044/2019_JSLHR-19-00201

Bennett, Meyer, Eikelboom, & Atlas. (2018). Investigating the Knowledge,

Skills, and Tasks Required for Hearing Aid Management: Perspectives of Clinicians and Hearing Aid Owners. *American Journal of Audiology*, *27*(1), 67–84. https://pubmed.ncbi.nlm.nih.gov/29305614/

Bennett, R. J., Saulsman, L., Eikelboom, R. H., & Olaithe, M. (2022). Coping with the social challenges and emotional distress associated with hearing loss: a qualitative investigation using Leventhal's self-regulation theory. *International Journal of Audiology*, *61*(5), 353–364. https://doi.org/10.1080/14992027.2021.1933620

Benus, S., Gravano, A., & Hirschberg, J. (2007). The Prosody of Backchannels in American English. *Proceedings of the 16th International Congress of Phonetic Sciences*, *August*, 1065–1068.

Berry, D. S., & Hansen, J. S. (1996). Positive Affect, Negative Affect, and Social Interaction. *Journal of Personality and Social Psychology*, *71*(4), 796–809. https://doi.org/10.1037/0022-3514.71.4.796

Beyan, C., Capozzi, F., Becchio, C., & Murino, V. (2018). Prediction of the leadership style of an emergent leader using audio and visual nonverbal features. *IEEE Transactions on Multimedia*, *20*(2), 441–456. https://doi.org/10.1109/TMM.2017.2740062

Bischoping, K. (1993). Gender differences in conversation topics, 1922-1990. *Sex Roles*, *28*(1–2), 1–18. https://doi.org/10.1007/BF00289744

Boholm, M., & Allwood, J. (2010). Repeated head movements, their function and relation to speech. *Proceedings of the Workshop on Multimodal Corpora Advances in Capturing Coding and Analyzing Multimodality LREC 2010*, 6–10. http://sskkii.gu.se/jens/publications/bfiles/B93.pdf

Bottalico, P., Piper, R. N., & Legner, B. (2022). Lombard effect, intelligibility, ambient noise, and willingness to spend time and money in a restaurant amongst older adults. *Scientific Reports*, *12*(1), 1–9. https://doi.org/10.1038/s41598-022-10414-6

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brimijoin, W. O., & Akeroyd, M. A. (2012). The role of head movements and

signal spectrum in an auditory front/back illusion. *I-Perception*, *3*(3), 179–181. https://doi.org/10.1068/i7173sas

Brimijoin, W. O., McShefferty, D., & Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *The Journal of the Acoustical Society of America*, *127*(6), 3678–3688. https://doi.org/10.1121/1.3409488

Brooks, D. N., Hallam, R. S., & Mellor, P. A. (2001). The effects on significant others of providing a hearing aid to the hearing-impaired partner. *British Journal of Audiology*, *35*(3), 165–171. https://doi.org/10.1080/00305364.2001.11745234

Brumm, H., & Zollinger, A. (2011). The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour*, *148*(11–13), 1173–1198. https://doi.org/10.1163/000579511X605759

Burke, L. A., & Naylor, G. (2020). Daily-Life Fatigue in Mild to Moderate Hearing Impairment: An Ecological Momentary Assessment Study. *Ear and Hearing*, *41*(6), 1518–1532. https://doi.org/10.1097/AUD.0000000000000888

Burns, M. E., & Pearson, J. C. (2011). An exploration of family communication environment, everyday talk, and family satisfaction. *Communication Studies*, *62*(2), 171–185. https://doi.org/10.1080/10510974.2010.523507

Carlile, S., & Keidser, G. (2020). Conversational Interaction Is the Brain in Action: Implications for the Evaluation of Hearing and Hearing Interventions. *Ear and Hearing*, *41*, 56S-67S. https://doi.org/10.1097/AUD.0000000000000939

Chartrand, T. L., & Bargh, J. A. (1999). 1999_the_Chameleon_Effect.Pdf. In *Journal of Personality and Social Psychology* (Vol. 76, Issue 6, pp. 893–910).

Ciorba, A., Bianchini, C., Pelucchi, S., & Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical Interventions in Aging*, *7*, 159–163. https://doi.org/10.2147/CIA.S26059

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and

reference. In *Advances in Psychology* (Vol. 9, Issue C). North-Holland Publishing Company. https://doi.org/10.1016/S0166-4115(09)60059-5

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process, Cognition, 22 (1986) l-39 1. *Cognition*, *22*, 1–39. https://pdfs.semanticscholar.org/dd2b/dd2c4df589cc3be1f4bfab6c42d8a9dc6609.pdf

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V., & Cowan, B. R. (2019). What makes a good conversation? Challenges in designing truly conversational agents. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3290605.3300705

Cooney, G., Mastroianni, A. M., Abi-Esber, N., & Brooks, A. W. (2020). The many minds problem: disclosure in dyadic versus group conversation. *Current Opinion in Psychology*, *31*, 22–27. https://doi.org/10.1016/j.copsyc.2019.06.032

Cunningham, D. W., Kleiner, M., Bülthoff, H. H., & Wallraven, C. (2004). The components of conversational facial expressions. *Proceedings - 1st Symposium on Applied Perception in Graphics and Visualization, APGV 2004*, *1*(212), 143–149. https://doi.org/10.1145/1012551.1012578

Cuperman, R., & Ickes, W. (2009). Big Five Predictors of Behavior and Perceptions in Initial Dyadic Interactions: Personality Similarity Helps Extraverts and Introverts, but Hurts "Disagreeables." *Journal of Personality and Social Psychology*, *97*(4), 667–684. https://doi.org/10.1037/a0015741

Dalton, D. S., Cruickshanks, K. J., Klein, B. E. K., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The Impact of Hearing Loss on Quality of Life in Older Adults. *Gerontologist*, *43*(5), 661–668. https://doi.org/10.1093/geront/43.5.661

Danescu-Niculescu-mizil, C., Gamon, M., & Dumais, S. (2011). Mark my words! Linguistic style accommodation in social media. *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*, 745–

754. https://doi.org/10.1145/1963405.1963509

Davis, M. H. (1980). Interpersonal Reactivity Index (IRI). *APA PsycTests*. https://doi.org/10.1037/t01093-000

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1–40. https://doi.org/10.1016/s0065-2601(01)80003-4

Drijvers, L., & Holler, J. (2022). Face-to-face spatial orientation fine-tunes the brain for neurocognitive processing in conversation. *IScience*, *25*(11), 105413. https://doi.org/10.1016/j.isci.2022.105413

Drimalla, H., Landwehr, N., Hess, U., & Dziobek, I. (2019). From face to face: the contribution of facial mimicry to cognitive and emotional empathy. *Cognition and Emotion*, *33*(8), 1672–1686. https://doi.org/10.1080/02699931.2019.1596068

Dunbar, R. I. M., Duncan, N. D. C., & Nettle, D. (1995). Size and structure of freely forming conversational groups. *Human Nature*, *6*(1), 67–78. https://doi.org/10.1007/BF02734136

Eagle, N., & Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, *10*(4), 255–268. https://doi.org/10.1007/s00779-005-0046-3

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). *Shared Reality Experiencing Commonality With Others' Inner States About the World*. *4*(5), 496–521.

Epley, N., & Schroeder, J. (2014). Mistakenly seeking solitude. *Journal of Experimental Psychology: General*, *143*(5), 1980–1999. https://doi.org/10.1037/a0037323

Fay, N., Garrod, S., & Carletta, J. (2000). Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science*, *11*(6), 481–486. https://doi.org/10.1111/1467-

9280.00292

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1477), 1669–1675. https://doi.org/10.1098/rspb.2001.1704

Floyd, K., Mikkelson, A. C., Tafoya, M. A., Farinelli, L., La Valley, A. G., Judd, J., Haynes, M. T., Davis, K. L., & Wilson, J. (2007). Human affection exchange: XIII. Affectionate communication accelerates neuroendocrine stress recovery. *Health Communication*, *22*(2), 123–132. https://doi.org/10.1080/10410230701454015

Foltz, A., Gaspers, J., Meyer, C., Thiele, K., Cimiano, P., & Stenneken, P. (2015). Temporal Effects of Alignment in Text-Based, Task-Oriented Discourse. *Discourse Processes*, *52*(8), 609–641. https://doi.org/10.1080/0163853X.2014.977696

Frick, R. W. (1985). Communicating Emotion : The Role of Prosodic Features. *Psychological Bulletin*, *97*(3), 412–429.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science*, *23*(8), 931–939. https://doi.org/10.1177/0956797612436816

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: A review. *Frontiers in Human Neuroscience*, *5*(FEBRUARY), 1–15. https://doi.org/10.3389/fnhum.2011.00011

Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes and Intergroup Relations*, *8*(2 SPEC. ISS.), 109–124. https://doi.org/10.1177/1368430205051060

Garnier, M., Henrich, N., & Dubois, D. (2010). Influence of sound immersion and communicative interaction on the lombard effect. *Journal of Speech, Language, and Hearing Research*, *53*(3), 588–608.

https://doi.org/10.1044/1092-4388(2009/08-0138)

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Gatica-Perez, D., McCowan, L., Zhang, D., & Bengio, S. (2005). Detecting group interest-level in meetings. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *I*(February 2005). https://doi.org/10.1109/ICASSP.2005.1415157

Glen, P. (2003). Laughter in Interaction. In *Cambridge: Cambridge University Press*.

Glista, D., O'Hagan, R., Moodie, S., & Scollie, S. (2020). An examination of clinical uptake factors for remote hearing aid support: a concept mapping study with audiologists. *International Journal of Audiology*, *60*(S1), S13–S22. https://doi.org/10.1080/14992027.2020.1795281

Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, *19*, 283–307.

Granberg, S., & Gustafsson, J. (2021). Key findings about hearing loss in the working-life: a scoping review from a well-being perspective. *International Journal of Audiology*, *60*(S2), 60–70. https://doi.org/10.1080/14992027.2021.1881628

Grebelsky-Lichtman, T. (2017). Verbal versus nonverbal primacy: Children's response to parental incongruent communication. *Journal of Social and Personal Relationships*, *34*(5), 636–661. https://doi.org/10.1177/0265407516651158

Grebelsky-Lichtman, T. (2021). Discrepant Verbal–Nonverbal Profile Theory. In *Engaging Theories in Interpersonal Communication* (3rd ed.). Routledge.

Grice, H. P. (1975). Logic and Conversation. *In Syntax and Semantics. Leiden, The Netherlands: Brill. Doi: Https://Doi.Org/10.1163/9789004368811_003*.

Guydish, A. J., & Fox Tree, J. E. (2023). In Pursuit of a Good Conversation: How Contribution Balance, Common Ground, and Conversational Closings Influence Conversation Assessment and Conversational

Memory. *Discourse Processes*, *60*(1), 18–41. https://doi.org/10.1080/0163853X.2022.2152552

Hadar, U., Steiner, T. J., Grant, E. C., & Clifford Rose, F. (1983). Head movement correlates of juncture and stress at sentence level. *Language and Speech*, *26*(2), 117–129. https://doi.org/10.1177/002383098302600202

Hadley, Fisher, & Pickering. (2020). Listeners are better at predicting speakers similar to themselves. *Acta Psychologica*, *208*. https://doi.org/10.1016/J.ACTPSY.2020.103094

Hadley, L. V., & Culling, J. F. (2022). Timing of head turns to upcoming talkers in triadic conversation: Evidence for prediction of turn ends and interruptions. *Frontiers in Psychology*, *13*(December), 1–14. https://doi.org/10.3389/fpsyg.2022.1061582

Hadley, L. V., Naylor, G., & Hamilton, A. F. de C. (2022). A review of theories and methods in the science of face-to-face social interaction. *Nature Reviews Psychology*, *1*(1), 42–54. https://doi.org/10.1038/S44159-021-00008-W

Hadley, L. V., Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2021a). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. *Psychonomic Bulletin and Review*, *28*(2), 632–640. https://doi.org/10.3758/s13423-020-01821-9

Hadley, L. V, Brimijoin, W. O., & Whitmer, W. M. (2019). Speech, movement, and gaze behaviours during dyadic conversation in noise. *Scientific Reports*, *9*(1), 1–8.

Hadley, L. V, Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2021b). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. *Psychonomic Bulletin & Review*, *28*(2), 632–640.

Hall, E. T. (1966). The Hidden Dimension. In *Anchor Books Editions* (Reprint. O, Issue 1).

Hall, J. A., & Davis, D. C. (2017). Proposing the Communicate Bond Belong Theory: Evolutionary Intersections With Episodic Interpersonal

Communication. *Communication Theory*, *27*(1), 21–47. https://doi.org/10.1111/comt.12106

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159–177. https://doi.org/10.1007/s40167-016-0044-5

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences*, *18*, 83–90. https://doi.org/10.1016/j.cobeha.2017.07.018

Harmsen, I. E. (2019). Empathy in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(10), 3939–3955. https://doi.org/10.1007/s10803-019-04087-w

Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, *130*(4), 2139–2152. https://doi.org/10.1121/1.3623753

Healy, E. W., & Yoho, S. E. (2016). Difficulty understanding speech in noise by the hearing impaired: Underlying causes and technological solutions. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2016-Octob*, 89–92. https://doi.org/10.1109/EMBC.2016.7590647

Hecht, M. L. (1978). the Conceptualization and Measurement of Interpersonal Communication Satisfaction. *Human Communication Research*, *4*(3), 253–264. https://doi.org/10.1111/j.1468-2958.1978.tb00614.x

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. https://doi.org/10.1016/j.wocn.2010.08.002

Helvik, A. S., Krokstad, S., & Tambs, K. (2013). Hearing loss and risk of early retirement. The HUNT study. *European Journal of Public Health*, *23*(4), 617–622. https://doi.org/10.1093/eurpub/cks118

Hepburn, A., Wilkinson, S., & Shaw, R. (2012). Repairing Self- and Recipient Reference. *Research on Language and Social Interaction*, *45*(2), 175–190.

https://doi.org/10.1080/08351813.2012.673914

Hill, C. E., Thompson, B. J., & Ladany, N. (2003). Therapist use of silence in therapy: A survey. *Journal of Clinical Psychology*, *59*(4), 513–524. https://doi.org/10.1002/jclp.10155

Hirschi, Q., Wilson, T. D., & Gilbert, D. T. (2023). Speak Up! Mistaken Beliefs About How Much to Talk in Conversations. *Personality and Social Psychology Bulletin*, *49*(10), 1454–1465. https://doi.org/10.1177/01461672221104927

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma, D., & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-short). *Journal of Autism and Developmental Disorders*, *41*(5), 589–596. https://doi.org/10.1007/s10803-010-1073-0

Hogan, A., O'Loughlin, K., Davis, A., & Kendig, H. (2009). Hearing loss and paid employment: Australian population survey findings. *International Journal of Audiology*, *48*(3), 117–122. https://doi.org/10.1080/14992020802449008

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006

Holler, J., & Wilkin, K. (2011). *Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue*. https://doi.org/10.1007/s10919-011-0105-6

Holman, J. A., Ali, Y. H. K., & Naylor, G. (2023). A qualitative investigation of the hearing and hearing-aid related emotional states experienced by adults with hearing loss. *International Journal of Audiology*, *62*(10), 973–982. https://doi.org/10.1080/14992027.2022.2111373

Holman, J. A., Drummond, A., Hughes, S. E., & Naylor, G. (2019). Hearing impairment and daily-life fatigue: a qualitative study. *International Journal of Audiology*, *58*(7), 408–416. https://doi.org/10.1080/14992027.2019.1597284

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

Huang, K., Yeomans, M., Brooks, A. W., Minson, J., & Gino, F. (2017). It doesn't hurt to ask: Question-asking increases liking. *Journal of Personality and Social Psychology*, *113*(3), 430–452. https://doi.org/10.1037/pspi0000097

Hygge, S., Ronnberg, J., Larsby, B., & Arlinger, S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *Journal of Speech and Hearing Research*, *35*(1), 208–215. https://doi.org/10.1044/jshr.3501.208

IEEE Recommended Practice for Speech Quality Measurements. (1969). *IEEE Transactions on Audio and Electroacoustics*, *17*(3), 225–246. https://doi.org/10.1109/IEEESTD.1969.7405210

Intille, S., Haynes, C., Maniar, D., Ponnada, A., & Manjourides, J. (2016). μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1124–1135. https://doi.org/10.1145/2971648.2971717

Ishi, C. T., Ishiguro, H., & Hagita, N. (2014). Analysis of relationship between head motion events and speech in dialogue conversations. *Speech Communication*, *57*, 233–243. https://doi.org/10.1016/j.specom.2013.06.008

ISO. (2003). *Ergonomic assessment of speech communication*. *Geneva*(9921–1).

Jakobson, R. (1972). Motor signs for 'Yes' and 'No.' *Language in Society*, *1*(1), 91–96. https://doi.org/10.1017/S0047404500006564

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a "standard maximum" silence of approximately one second in conversation. In *Conversation: An interdisciplinary pespective* (pp. 166–196).

Jensen, M. (2016). Personality traits and nonverbal communication patterns. *International Journal of Social Science Studies*, *4*(5). https://doi.org/10.11114/ijsss.v4i5.1451

Joore, M. A., Potjewijd, J., Timmerman, A. A., & Anteunis, L. J. C. (2002). Response shift in the measurement of quality of life in hearing impaired adults after hearing aid fitting. *Quality of Life Research*, *11*(4), 299–307. https://doi.org/10.1023/A:1015598807510

Kagan, A., Winckel, J., Black, S., Duchan, J. F., Simmons-Mackie, N., & Square, P. (2004). A Set of Observational Measures for Rating Support and Participation in Conversation between Adults with Aphasia and Their Conversation Partners. *Topics in Stroke Rehabilitation*, *11*(1), 67–83. https://doi.org/10.1310/CL3V-A94A-DE5C-CVBE

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, *146*(4), 569–576. https://doi.org/10.1037/xge0000271

Keidser, G., Naylor, G., Brungart, D. S., Caduff, A., Campos, J., Carlile, S., Carpenter, M. G., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., & Smeds, K. (2020). The Quest for Ecological Validity in Hearing Science: What It Is, Why It Matters, and How to Advance It. *Ear and Hearing*, *41*, 5S-19S. https://doi.org/10.1097/AUD.0000000000000944

Kemper, S. (1994). Elderspeak: Speech Accommodations to Older Adults. *Aging, Neuropsychology, and Cognition*, *1*(1), 17–28. https://doi.org/10.1080/09289919408251447

Kendon, A. (2002). Some uses of the head shake. *Gesture*, *2*(2), 147–182. https://doi.org/10.1075/gest.2.2.03ken

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., Davis, A. C., Erber, N. P., Hickson, L., Holmes, A., Rosenhall, U., & Von Wedel, H. (2003). Candidature for and delivery of audiological services: Special needs of older people. *International Journal of Audiology*, *42*(SUPPL. 2). https://doi.org/10.3109/14992020309074650

Kohler, K. J. (2008). “Speech-smile”, “speech-laugh”, “laughter” and their sequencing in dialogic interaction. *Phonetica*, *65*(1–2), 1–18. https://doi.org/10.1159/000130013

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2011). Disrupting the flow: How brief silences in group conversations affect social needs. *Journal of Experimental Social Psychology*, *47*(2), 512–515. https://doi.org/10.1016/j.jesp.2010.12.006

Koutsombogera, M., & Vogel, C. (2022). Understanding Laughter in Dialog. *Cognitive Computation*, *14*(4), 1405–1420. https://doi.org/10.1007/s12559-022-10013-7

Krepsz, V., Horváth, V., Huszár, A., Neuberger, T., & Gyarmathy, D. (2024). ‘Should we laugh?’ Acoustic features of (in)voluntary laughters in spontaneous conversations. *Cognitive Processing*, *25*(1), 89–106. https://doi.org/10.1007/s10339-023-01168-8

Kumano, S., Ishii, R., & Otsuka, K. (2017). Comparing empathy perceived by interlocutors in multiparty conversation and external observers. *2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, *2018-Janua*, 50–57. https://doi.org/10.1109/ACII.2017.8273578

Lavan, N., Scott, S. K., & McGettigan, C. (2016). Laugh Like You Mean It: Authenticity Modulates Acoustic, Physiological and Perceptual Properties of Laughter. *Journal of Nonverbal Behavior*, *40*(2), 133–149. https://doi.org/10.1007/s10919-015-0222-8

Lee, C. C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., Georgiou, P. G., & Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. *Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010*, *January*, 793–796. https://doi.org/10.21437/interspeech.2010-287

Leertouwer, I., Schuurman, N. K., & Vermunt, J. K. (2022). Are Retrospective Assessments Means of People’s Experiences? Accounting for Interpersonal and Intrapersonal Variability when Comparing

Retrospective Assessment Data to Ecological Momentary Assessment Data. *Journal for Person-Oriented Research*, *8*(2), 52–70. https://doi.org/10.17505/jpor.2022.24855

Lelic, D., Wolters, F., & Schinkel-Bielefeld, N. (2023). Journal of the American Academy of Audiology Measuring hearing aid satisfaction in everyday listening situations: Retrospective and in-situ assessments complement each other. *Journal of the American Academy of Audiology*. https://doi.org/10.1055/a-2265-9418

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). *Package "Emmeans" R Package Version 4.0-3*. http://cran.r-project.org/package=emmeans

Levinson, S. C. (2020). On the human "interaction engine." *Roots of Human Sociality: Culture, Cognition and Interaction*, 39–69. https://doi.org/10.4324/9781003135517-3

Levinson, S. C., Holler, J., & Levinson, S. C. (2014). *The origin of human multi-modal communication*.

Levitan, R., Gravano, A., Willson, L., Beňuš, Š., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. *NAACL HLT 2012 - 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, 11–19.

Li, C. M., Zhang, X., Hoffman, H. J., Cotch, M. F., Themann, C. L., & Wilson, M. R. (2014). Hearing impairment associated with depression in US adults, National Health and Nutrition Examination Survey 2005-2010. *JAMA Otolaryngology - Head and Neck Surgery*, *140*(4), 293–302. https://doi.org/10.1001/jamaoto.2014.42

Lind, C. (2013). Conversation repair: Ecological validity of outcome measures in acquired hearing impairment. *Cochlear Implants International*, *14*(SUPPL. 4). https://doi.org/10.1179/1467010013Z.000000000135

Lind, C., Hickson, L., & Erber, N. (2010). Who said what? Sampling conversation repair behavior involving adults with acquired hearing impairment. *Seminars in Hearing*, *31*(2), 104–115.

https://doi.org/10.1055/s-0030-1252104

Lind, C., Hickson, L., & Erber, N. P. (2004). Conversation repair and acquired hearing impairment: A preliminary quantitative clinical study. *Australian and New Zealand Journal of Audiology*, *26*(1), 40–52. https://doi.org/10.1375/audi.26.1.40.55987

Lind, C., Hickson, L., & Erber, N. P. (2006). Conversation repair and adult cochlear implantation: A qualitative case study. *Cochlear Implants International*, *7*(1), 33–48. https://doi.org/10.1002/cii.20

Littlejohn and Foss. (2011). *Theories of human communication - communication theory and scholarship.pdf*.

Liu, M., Duan, Y., Ince, R. A. A., Chen, C., Garrod, O. G. B., Schyns, P. G., & Jack, R. E. (2022). Facial expressions elicit multiplexed perceptions of emotion categories and dimensions. *Current Biology*, *32*(1), 200-209.e6. https://doi.org/10.1016/j.cub.2021.10.035

Louven, C., Scholle, C., Gehrs, F., & Lenz, A. (2022). emoTouch Web – a web-based system for continuous real time studies with smartphones, tablets, and desktop computers. *Jahrbuch Musikpsychologie*, *30*(April). https://doi.org/10.5964/jbdgm.137

Manley, G. A. (2017). Comparative Auditory Neuroscience: Understanding the Evolution and Function of Ears. *JARO - Journal of the Association for Research in Otolaryngology*, *18*(1), 1–24. https://doi.org/10.1007/s10162-016-0579-3

Manning, J. (2014). A Constitutive Approach to Interpersonal Communication Studies. *Communication Studies*, *65*(4), 432–440. https://doi.org/10.1080/10510974.2014.927294

Maynard, S. K. (1987). Interactional functions of a nonverbal sign Head movement in japanese dyadic casual conversation. *Journal of Pragmatics*, *11*(5), 589–606. https://doi.org/10.1016/0378-2166(87)90181-0

McCaffrey, S. A., Chiauzzi, E., Chan, C., & Hoole, M. (2018). Understanding 'Good Health care' from the Patient's Perspective: Development of a Conceptual Model Using Group Concept Mapping. *The Patient - Patient-Centered Outcomes Research 2018 12:1*, *12*(1), 83–95.

https://doi.org/10.1007/S40271-018-0320-X

McFarland, D. A., Jurafsky, D., & Rawlings, C. (2013). Making the connection: Social bonding in courtship situations. *American Journal of Sociology*, *118*(6), 1596–1649. https://doi.org/10.1086/670240

McLaughlin, M. L., & Cody, M. J. (1982). Awkward Silences: Behavioral Antecedents and Consequences of the Conversational Lapse. *Human Communication Research*, *8*(4), 299–316. https://doi.org/10.1111/j.1468-2958.1982.tb00669.x

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, *49*(2), 201–213. https://doi.org/10.1016/S0749-596X(03)00028-7

Meyer, A. S. (2023). Timing in Conversation. *Journal of Cognition*, *6*(1), 1–17. https://doi.org/10.5334/joc.268

Meyer, C., Waite, M., Atkins, J., Ekberg, K., Scarinci, N., Barr, C., Cowan, R., & Hickson, L. (2021). How Can eHealth Meet the Hearing and Communication Needs of Adults With Hearing Impairment and their Significant Others? A Group Concept Mapping Study. *Ear & Hearing*, *Publish Ah*, 1–12. https://doi.org/10.1097/aud.0000000000001097

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, *20*(3), 336–366. https://doi.org/10.1111/josl.1_12177

Monnet, T., Samson, M., Bernard, A., David, L., & Lacouture, P. (2014). Measurement of three-dimensional hand kinematics during swimming with a motion capture system: A feasibility study. *Sports Engineering*, *17*(3), 171–181. https://doi.org/10.1007/s12283-014-0152-4

Moore, H. T. (1992). Further data concerning sex differences. *The Journal of Abnormal Psychology and Social Psychology*, *17(2)*, 210–214. https://doi.org/10.1037/h0064645

Morgan, A., Hickson, L., & Worrall, L. (2002). The impact of hearing impairment on quality of life of older people. *Asia Pacific Journal of Speech, Language and Hearing*, *7*(1), 39–53.

https://doi.org/10.1179/136132802805576562

Most, T., & Aviner, C. (2009). Auditory, visual, and auditory - Visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, *14*(4), 449–464. https://doi.org/10.1093/deafed/enp007

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory-visual perception of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research*, *55*(1148–1162).

Most, T., Weisel, A., & Zaychik, A. (1993). Auditory, visual and auditory—visual identification of emotions by hearing and hearing-impaired adolescents. *British Journal of Audiology*, *27*(4), 247–253.

Mulwa, K. W., & Kucker, S. C. (2024). Coding social interactions in naturalistic settings: The taxonomy of dyadic conversation. *Behavior Research Methods*, *56*(1), 172–186. https://doi.org/10.3758/s13428-022-02033-w

Murphy, N. A., & Hall, J. A. (2021). Capturing Behavior in Small Doses: A Review of Comparative Research in Evaluating Thin Slices for Behavioral Measurement. *Frontiers in Psychology*, *12*(April), 1–13. https://doi.org/10.3389/fpsyg.2021.667326

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 169–172. https://doi.org/10.3115/1557690.1557737

Niederehe, G., & Duncan, S. (1974). On Signalling That it's Your Turn to Speak. *Journal of Experimental Social Psychology*, *10*(3), 234―247.

Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International Journal of Audiology*, *52*(6), 409–412. https://doi.org/10.3109/14992027.2013.781278.A

Nölle, J., Chen, C., Hensel, L. B., Garrod, O. G. B., Schyns, P. G., & Jack, R. E. (2021). Facial expressions of emotion include iconic signals of

rejection and acceptance. *Journal of Vision*, *21*(9), 2932. https://doi.org/10.1167/jov.21.9.2932

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision*, *8*(8), 1–23. https://doi.org/10.1167/8.8.1

Oertel, C., Scherer, S., & Campbell, N. (2011). On the use of multimodal cues for the prediction of degrees of involvement in spontaneous conversation. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *January 2014*, 1541–1544. https://doi.org/10.21437/interspeech.2011-465

Okada, S., Ohtake, Y., Nakano, Y. I., Hayashi, Y., Huang, H. H., Takase, Y., & Nitta, K. (2016). Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets. *ICMI 2016 - Proceedings of the 18th ACM International Conference on Multimodal Interaction*, *1*, 169–176. https://doi.org/10.1145/2993148.2993154

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2011). Best Practices in Exploratory Factor Analysis. In *Best Practices in Quantitative Methods* (Issue October). https://doi.org/10.4135/9781412995627.d8

Ostrand, R., & Chodroff, E. (2021). It's alignment all the way down, but not all the way up: Speakers align on some features but not others within a dialogue. *Journal of Phonetics*, *88*, 101074. https://doi.org/10.1016/j.wocn.2021.101074

Otsuka, K., & Tsumori, M. (2020). Analyzing Multifunctionality of Head Movements in Face-to-Face Conversations Using Deep Convolutional Neural Networks. *IEEE Access*, *8*, 217169–217195. https://doi.org/10.1109/ACCESS.2020.3041672

Palmer, A. D., Carder, P. C., White, D. L., Saunders, G., Woo, H., Graville, D. J., & Newsom, J. T. (2019). The impact of communication impairments on the social relationships of older adults: Pathways to psychological well-being. *Journal of Speech, Language, and Hearing Research*, *62*(1), 1–21. https://doi.org/10.1044/2018_JSLHR-S-17-0495

Paluch, R., Latzel, M., Meis, M., & RICHARD PALUCH MATTHIAS LATZEL AND MARKUS MEIS. (2015). A new tool for subjective assessment of hearing aid performance: Analyses of Interpersonal Communication. *Proceedings of the International Symposium on Auditory and Audiological Research*, *5*(August), 453–460.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, *119*(4), 2382–2393. https://doi.org/10.1121/1.2178720

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2018). Similar neural responses predict friendship. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02722-7

Paxton, A., & Dale, R. (2017). Interpersonal movement synchrony responds to high- and low-level conversational constraints. *Frontiers in Psychology*, *8*(JUL), 1–16. https://doi.org/10.3389/fpsyg.2017.01135

Peperkoorn, L. S., Becker, D. V., Balliet, D., Columbus, S., Molho, C., & Van Lange, P. A. M. (2020). The prevalence of dyads in social life. *PLoS ONE*, *15*(12 December), 1–17. https://doi.org/10.1371/journal.pone.0244188

Petersen, E. B., MacDonald, E. N., & Josefine Munch Sørensen, A. (2022). The Effects of Hearing-Aid Amplification and Noise on Conversational Dynamics Between Normal-Hearing and Hearing-Impaired Talkers. *Trends in Hearing*, *26*. https://doi.org/10.1177/23312165221103340

Petersen, E., Walravens, El., & Kofoed Pedersen, A. (2023). Backchannel Behavior in Conversations and How it is Affected by Hearing Loss, Noise, and Hearing Aids. *10th Convention of the European Acoustics Association*, *September*, 2417–2424. https://doi.org/10.61782/fa.2023.0432

Picou, E. M. (2022). Hearing Aid Benefit and Satisfaction Results from the MarkeTrak 2022 Survey: Importance of Features and Hearing Care Professionals. *Seminars in Hearing*, *43*(4), 301–316. https://doi.org/10.1055/s-0042-1758375

Poost-Foroosh, L., Jennings, M. B., Shaw, L., Meston, C. N., & Cheesman, M.

F. (2011). Factors in Client–Clinician Interaction That Influence Hearing Aid Adoption. *Trends in Amplification*, *15*(3), 127–139. https://doi.org/10.1177/1084713811430217

Provine, R. . (1996). Contagious Yawning and Laughter: Significance for Sensory Feature Detection, Motor Pattern Generation, Imitation, and the Evolution of Social Behavior. In *Social Learning in Animals*. ACADEMIC PRESS, INC. https://doi.org/10.1016/b978-012273965-1/50010-8

Provine, R. R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, *30*(1), 1–4. https://doi.org/10.3758/BF03330380

Provine, R. R., & Fischer, K. R. (1989). Laughing, Smiling, and Talking: Relation to Sleeping and Social Context in Humans. *Ethology*, *83*(4), 295–305. https://doi.org/10.1111/j.1439-0310.1989.tb00536.x

Raman, C., Prabhu, N. R., & Hung, H. (2023). Perceived Conversation Quality in Spontaneous Interactions. *IEEE Transactions on Affective Computing*, *14*(4), 2901–2912. https://doi.org/10.1109/TAFFC.2023.3233950

Ramsberger, G., & Rende, B. (2002). Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, *16*(3), 337–353. https://doi.org/10.1080/02687040143000636

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2018). Daily well-being: The role of autonomy, competence, and relatedness. *Relationships, Well-Being and Behaviour: Selected Works of Harry Reis*, 317–349. https://doi.org/10.4324/9780203732496-13

Reitter, D., Keller, F., & Moore, J. D. (2011). A Computational Cognitive Model of Syntactic Priming. *Cognitive Science*, *35*(4), 587–637. https://doi.org/10.1111/j.1551-6709.2010.01165.x

Richmond, V. P., McCroskey, J. C., McCroskey, L. L., & Fayer, J. M. (2008). Communication Traits In First And Second Languages. *Journal of Intercultural Communication Research*, *37*(2), 65–73. https://doi.org/10.1080/17475750802533331

Riest, C., Jorschick, A. B., & de Ruiter, J. P. (2015). Anticipation in turn-

taking: Mechanisms and information sources. *Frontiers in Psychology*, *6*(FEB), 1–14. https://doi.org/10.3389/fpsyg.2015.00089

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247–261. https://doi.org/10.1080/14992027.2018.1551631

Rosas, S., & Kane, M. (2012). Quality and rigor of the concept mapping methodology: a pooled study analysis. *Evaluation and Program Planning*, *35*(2), 236–245. https://doi.org/10.1016/J.EVALPROGPLAN.2011.10.003

Rubin, R. B., Perse, E. M., & Barbato, C. A. (1988). Conceptualization and Measurement of Interpersonal Communication Motives I shape and nurture their self-concepts, make decisions about. *Human Communication Research International Communication Association*, *14*(4), 602–628.

Russell, D., Peplau, L., & Ferguson, M. L. (1978). UCLA Loneliness Scale. *APA PsycTests*. https://doi.org/10.1037/t05974-000

Sacks, H, Schegloff, E., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation Harvey Sacks , Emanuel A . Schegloff and Gail Jefferson. *Language*, *50*(4), 696–735.

Satterthwaite, F. E. (1946). An Approximate Distribution on Estimates of Variance Components. *Biometrics Bulletin*, *2*, 110–114. http://dx.doi.org/10.2307/3002019

Sauter, D. (2010). More than happy: The need for disentangling positive emotions. *Current Directions in Psychological Science*, *19*(1), 36–40. https://doi.org/10.1177/0963721409359290

Scarinci, N., Worrall, L., & Hickson, L. (2008). The effect of hearing impairment in older people on the spouse. *International Journal of Audiology*, *47*(3), 141–151. https://doi.org/10.1080/14992020701689696

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. *Roundtable on*

Languages and Linguistics, *March*, 71–93.

Schinkel-bielefeld, N. (2020). *Assessment : What Situations*. *29*(September), 591–609.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233. https://doi.org/10.1016/j.cognition.2009.08.009

See, A., Roller, S., Kiela, D., & Weston, J. (2019). What makes a good conversation? How controllable attributes affect human judgments. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 1702–1723. https://arxiv.org/abs/1902.08654v2

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shiell, M. M., Zaar, J., Høy-Christensen, J., Skoglund, M. A., Keidser, G., & Rotger-Griful, S. (2023). Multilevel Modeling of Gaze From Listeners With Hearing Loss Following a Realistic Conversation. *Journal of Speech, Language, and Hearing Research*, *66*(11), 4575–4589. https://doi.org/10.1044/2023_JSLHR-22-00641

Smeds, K., Gotowiec, S., Wolters, F., Herrlin, P., Larsson, J., & Dahlquist, M. (2020). *Smeds - Selecting_Scenarios_for_Hearing_Related_Laboratory.3-2*. 20–30.

Sørensen, A. J. M., Fereczkowski, M., & MacDonald, E. N. (2021). Effects of Noise and Second Language on Conversational Dynamics in Task Dialogue. *Trends in Hearing*, *25*. https://doi.org/10.1177/23312165211024482

Spitzberg, B. H., & Canary, D. J. (1985). *Loneliness and relationally competent communication* (pp. 387–402). Journal of Social and Personal relationships, Vol. 2.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, *25*(5–6), 701–721. https://doi.org/10.1023/a:1020867916902

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society*, *35*(3), 367–392. https://doi.org/10.1017/S0047404506060179

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., & Kaplan, G. A. (2000). Negative Consequences of Hearing Impairment in Old AgeA Longitudinal Analysis. *The Gerontologist*, *40*(3), 320–326. https://doi.org/10.1093/GERONT/40.3.320

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *Journal of the Acoustical Society of America*, *26*(2), 212–215. https://doi.org/10.1121/1.1907309

Tanaka, H., & Campbell, N. (2011). Acoustic Features of Four Types of Laughter in Natural Conversational Speech. *International Congress of Phonetic Sciences, August*. https://api.semanticscholar.org/CorpusID:22893960

Templeton, E. M., Chang, L. J., Reynolds, E. A., Cone Lebeaumont, M. D., & Wheatley, T. (2023). Long gaps between turns are awkward for strangers but not for friends. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1875). https://doi.org/10.1098/rstb.2021.0471

ter Bekke, M., Drijvers, L., & Holler, J. (2024). Hand Gestures Have Predictive Potential During Conversation: An Investigation of the Timing of Gestures in Relation to Speech. *Cognitive Science*, *48*(1). https://doi.org/10.1111/cogs.13407

To, W. M., & Chung, A. W. L. (2015). Restaurant noise: Levels and temporal characteristics. *Noise and Vibration Worldwide*, *46*(8), 11–17. https://doi.org/10.1260/0957-4565.46.8.11

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12*(1), 1–16. https://doi.org/10.1016/0149-7189(89)90016-5

Trochim, W., Cook, J. A., & Setze, R. J. (1994). Using Concept Mapping to Develop a Conceptual Framework of Staff's Views of a Supported Employment Program for Individuals With Severe Mental Illness. *Journal of Consulting and Clinical Psychology*, *62*(4), 766–775. https://doi.org/10.1037/0022-006X.62.4.766

Trujillo, J., Özyürek, A., Holler, J., & Drijvers, L. (2021). Speakers exhibit a multimodal Lombard effect in noise. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-95791-0

Trujillo, J. P., & Holler, J. (2024). Conversational facial signals combine into compositional meanings that change the interpretation of speaker intentions. *Scientific Reports*, *14*(1), 1–14. https://doi.org/10.1038/s41598-024-52589-0

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*(NOV). https://doi.org/10.3389/FPSYG.2014.01323/ABSTRACT

Tye-Murray, N. (2003). Conversational fluency of children who use cochlear implants. *Ear and Hearing*, *24*(1 SUPPL.), 82–89. https://doi.org/10.1097/01.aud.0000051691.33869.ec

Vas, V., Akeroyd, M. A., & Hall, D. A. (2017). *A Data-Driven Synthesis of Research Evidence for Domains of Hearing Loss , as Reported by Adults With Hearing Loss and Their Communication Partners*. *21*, 1–25. https://doi.org/10.1177/2331216517734088

Ventry, I. M., & Weinstein, B. E. (1982). The hearing handicap inventory for the elderly: A new tool. *Ear and Hearing*, *3*(3), 128–134. https://doi.org/10.1097/00003446-198205000-00006

Vettin, J., & Todt, D. (2004). Laughter in conversation: Features of occurrence

and acoustic structure. *Journal of Nonverbal Behavior*, *28*(2), 93–115. https://doi.org/10.1023/B:JONB.0000023654.73558.72

Vivek, Patil H., Singh, S. N., Mishra, S., & Donavan, D. T. (2007). *Parallel Analysis Engine to Aid Determining Number of Factors to Retain*. https://analytics.gonzaga.edu/parallelengine/

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Dilegge, T., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Fleming, T., Forouzanfar, M. H., … Zuhlke, L. J. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, *57*, 209–232. https://doi.org/10.1016/j.specom.2013.09.008

Warner-Garcia, S. (2014). Laughing when nothing's funny: The pragmatic use of coping laughter in the negotiation of conversational disagreement. *Pragmatics*, *24*(1), 157–180. https://doi.org/10.1075/prag.24.1.07war

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weick, K. E. (1979). *The Social Psychology of Organizing* (Second). McGraw-Hill (1979).

Whitmer, W. M., McShefferty, D., & Smeds, K. (2024). Exploring attentive listening in noise through the just-follow conversation task. *SPIN2024*, P04.

Wilson, J., Hickson, L., & Worrall, L. (1998). Use of communication strategies by adults with hearing impairment. *Asia Pacific Journal of Speech, Language and Hearing*, *3*(1), 29–42. https://doi.org/10.1179/136132898805577250

Wolters, F., Smeds, K., Schmidt, E., Christensen, E. K., & Norup, C. (2016). Common sound scenarios: A context-driven categorization of everyday sound environments for application in hearing-device research. *Journal of the American Academy of Audiology*, *27*(7), 527–540. https://doi.org/10.3766/jaaa.15105

World Health Organization. (2001). *ICF: International Classification of Functioning, Disability and Health*.

Wu, Y. H., Stangl, E., Chipara, O., Gudjonsdottir, A., Oleson, J., & Bentler, R. (2020). Comparison of In-Situ and Retrospective Self-Reports on Assessing Hearing Aid Outcomes. *Journal of the American Academy of Audiology*, *31*(10), 746–757. https://doi.org/10.1055/s-0040-1719133

Wynn, C. J., Barrett, T. S., & Borrie, S. A. (2022). Rhythm Perception, Speaking Rate Entrainment, and Conversational Quality: A Mediated Model. *Journal of Speech, Language, and Hearing Research*, *65*(6), 2187–2203. https://doi.org/10.1044/2022_JSLHR-21-00293

Yeomans, M., Schweitzer, M. E., & Brooks, A. W. (2022). The Conversational Circumplex: Identifying, prioritizing, and pursuing informational and relational motives in conversation. *Current Opinion in Psychology*, *44*, 293–302. https://doi.org/10.1016/j.copsyc.2021.10.001

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex*, *27*(10), 4988–5000. https://doi.org/10.1093/cercor/bhx202

Zevon, M. A., Schwabish, S., Donnelly, J. P., & Kerry, R. (2007). Medically related legal needs and quality of life in cancer care: a structural analysis. *Cancer*, *109*(12), 2600–2606. https://doi.org/10.1002/CNCR.22682