37(1)
April 2025

Parallèles

FACULTÉ DE TRADUCTION
ET D'INTERPRÉTATION

UNIVERSITÉ
DE GENÈVE

Parallèles 37(1) April 2025

Sommaire – Contents

 DOI:10.17462/para.2025.01.00

Parallèles 37(1) April 2025

Gesture in interpreting

Sílvia Gabarró-López
Université de Namur & Universitat Pompeu Fabra

Alan Cienki
Vrije Universiteit Amsterdam

Guest Editors

1. Gesture studies and interpreting studies: Long-time partners, and now a newlywed couple

The use of bodily articulators when humans communicate is a shared feature across different modalities, i.e., spoken, signed, or tactile. Hearing speakers, deaf-sighted signers, and deafblind signers use their bodily articulators to convey meaning in different ways depending on the language. The main articulators used in the signed and the tactile modalities are the two hands, which articulate various types of signs and gestures. Deaf-sighted signers also use facial movements and expressions as an integral part of visual sign languages. Although the main articulators used in spoken languages are found in the vocal tract, hearing speakers also produce manual and nonmanual gestures accompanying speech.

The integration of speech and gesture occurs naturally, so it can be observed not only in conversations among people sharing the same language but also in interpreted discourse. Modern gesture studies started in the 1970s with the works of Kendon (1972) and McNeill (1979). Since then, the field has flourished and approached gestures from different perspectives (e.g., pragmatics, L1 and L2 acquisition, gesture-sign integration, etc.), using different theories (e.g., those of Kendon, McNeill, Müller, etc.), methodologies (e.g., observational, corpus-based, or experimental studies), and terminology. These studies have been carried out mostly by linguists, psychologists, or anthropologists who have focused on non-interpreted discourse in different cultures around the world.

Interpreting studies started more or less at the same time as gesture studies, focusing on conference interpreting (Gile, 2009). Other interpreting settings (e.g., dialogue interpreting, sight translation, etc.) and theories (notably from psychology and linguistics) were progressively added, including both spoken and signed languages. Multimodality, understood here as the use of multiple bodily articulators, has been present from the onset of interpreting studies to a greater or lesser extent. When analyzing the renditions of sign language interpreters, scholars have considered the diverse semiotic resources employed by interpreters to convey meaning (e.g., Janzen & Shaffer, 2013; Nicodemus et al., 2017; Nilsson, 2015), mostly in the target signed language discourse. This tendency aligns with the fact that the spoken-to-sign language interpreting direction has been much more investigated than the signed-to-spoken language direction (Wang, 2021).

In spoken language interpreting, some scholars attested to the importance of resources beyond the verbal dimension in interpreting almost from the onset of the field (e.g., Anderson, 1979; Lang, 1976, 1978). The multimodal analysis of interpreted renditions followed in the coming decades and increased in the 2000s, but remained scattered, including a variety of settings (e.g., medical, legal, and pedagogical), embodied resources (e.g., nonmanual elements, seating

 DOI:10.17462/para.2025.01.01

arrangement, or manual activity), phenomena under scrutiny (e.g., dynamics of inclusion and exclusion of primary parties, and the coordinating role of the interpreter), and theoretical frameworks (e.g., Conversation Analysis and Discourse Analysis, among others) (see Davitti, 2019 for an overview).

As of today, the study of the gestural behavior of interpreters at work is mostly based on spoken-to-spoken language interpreting (e.g., Cienki & Iriskhanova, 2020; Zagar Galvão, 2020) and, to a lesser extent, on signed-to-spoken language interpreting (Bø, in press; Gabarró-López, 2024). To join efforts and allow scholars working on both modalities to speak to each other, we organized a panel on "Gestures in spoken-to-spoken and signed-to-spoken language interpreting" for the 18th conference of the *International Pragmatics Association* (IPrA) held in Brussels on 9–14 July 2023. To the best of our knowledge, this is the first volume devoted to this young emerging interdisciplinary field of inquiry.

2. Outline of the volume

This thematic issue contains nine selected papers from the IPrA panel covering different topics and language combinations (see Table 1). Most contributions analyze spoken-to-spoken language interpreting, and two focus on signed-to-spoken language interpreting[1]. We grouped the papers according to the setting in which interpreting took place and the methodology for data collection. The first four papers draw on data from interpreters working in a booth, whereas the other five papers worked on the renditions produced in other situations. All interpreters of the first type of papers interpreted monological speech simultaneously, mostly from recorded TED talks (Cienki, Martín de León & Zagar Galvão, and Ren & Wang) or from live discourse (Olza). Interpreters in the second type of papers also engaged in simultaneous interpreting of recorded dialogical or monological discourses (Gabarró-López and Janzen et al., respectively) or in live (simulated) dialogue interpreting (Beukeleers et al., Bø, and Chwalczuk).

Authors	Topics	Language combinations
Martín de León and Zagar Galvão	Comparative study of gestures in non-/interpreted discourse	English > Italian/Portuguese
Cienki	Gestures in (dis)fluent discourse	Russian > English/German
Ren and Wang	Gestures in disfluent discourse and cognitive load	English > Mandarin Chinese
Olza	Gestural alignment	Spanish > English/French
Janzen et al.	Gestural-conceptual alignment	English > French/Navajo/Spanish/ Ukrainian
Chwalczuk	Gestural profiles	English > Bengali/Indonesian/ Panjabi/ Portuguese/Spanish English/French > Arabic/Czech/ Dutch/German/ Hungarian/Italian/ Mandarin
Beukeleers et al.	Gestural omissions and additions	Turkish > Flemish Dutch Russian > Flemish Dutch

[1] The exclusion of the spoken-to-signed direction was a deliberate decision when we organized the panel. The reason behind it is that we wanted target productions to be similar, to foster discussion among participants: If we had included spoken-to-signed interpreting, the target discourses would have been too different and may have included other aspects that were beyond the scope of the gathering.

Bø	Gaze and head gestures	NTS (Norwegian Sign Language) > Norwegian
Gabarró-López	Reformulation structures	LSFB (French Belgian Sign Language) > Belgian French

Table 1. Topics and language combination of the contributions

In terms of the methods of analyzing gestures, some of the studies (Beukeleers et al., Olza, Ren & Wang) employ the mixed form/function system popularized in McNeill (1992), using the categories of iconic, metaphoric, deictic, and beat gestures. Other studies here (Chwalczuk, Cienki) rely on a specifically functional system of categorization, espoused in works such as Müller (1998), Bressem et al. (2013), and Cienki (2013), using categories such as representational (concrete or abstract), pragmatic, deictic, and self-adaptor gestures. We can note that this difference reflects a debate on methods of analysis which is ongoing in the field of gesture studies.

The following provides a more detailed overview of the individual contributions.

Celia Martín de León and Elena Zagar Galvão conducted a comparative study of the gestures produced by five professional conference interpreters while interpreting a TED Talk from their B language (English) to their A language (Portuguese or Italian), simultaneously inside the booth and in a face-to-face interview with one of the authors. The analysis reveals that interpreters produced more gestures in the first (monological) condition, but the forms of the gestures were larger in the second (dialogical) condition. Regarding the rate of self-adapters in the two conditions, the authors do not find a shared pattern by all interpreters of the dataset and conclude that this category of gestures needs to be further explored.

Alan Cienki and colleagues investigated gestures during disfluent and fluent speech in simultaneous interpreting. Forty-nine people who trained or worked as professional interpreters participated in the study. They had to interpret educational lectures about biodiversity from Russian (the A language of all participants) to English/German (one of them the B language of the participants) or vice versa. The results showed that the most frequently used gestures were pragmatic ones or self-adapters, whereas representational gestures were produced to a lesser extent. The author explains this distribution as a result of the functions of gestures. While self-adapters and pragmatic gestures can help regulate stress and organize structure discourse, respectively, representational gestures result from deeper semantic processing and may not be produced because of time constraints.

Similarly, **Yuetao Ren and Jianhua Wang** studied gestures produced in moments of disfluency in simultaneous interpreting and related their functions to cognitive load. Thirteen master's students were recorded while interpreting two talks from English (their B language) to Mandarin Chinese (their A language). The results suggest that gesture and cognitive load are interconnected, as most gestures occurred with or adjacent to processing difficulties. Furthermore, different types of gestures were produced depending on the disfluency: beat and metaphoric gestures were often produced during silence, whereas deictic and iconic gestures were used less frequently. Interestingly, this latter finding regarding iconic/representational gestures is in line with Cienki (this volume).

Inés Olza explored the gestural alignment of two novice interpreters (interpreting into English or French) with the source speaker, who produces a monological talk on the history of technology in Spanish. The study finds alignment between the source speaker and the interpreters on the general level of gesture production; interpreters were more likely to gesture than not at moments when the source speaker gestured. However, in terms of the types of gesture used,

the findings suggest that gestural alignment may not necessarily be driven by the distinction between whether the gestures are related to the content of the speech or not. Instead, other factors of the source speaker's behavior may be more influential in determining when the interpreter gestures, such as the rhythm of the speech and other prosodic features.

Terry Janzen, Lorraine Leeson, and Barbara Shaffer also examined gestural alignment and combined it with conceptual alignment in their analysis, considering both of these between source text speakers and their interpreters. Their study considered 14 professional simultaneous interpreters; each interpreted talk in English from two different speakers into either French, Navajo, Spanish, or Ukrainian. The data reveal instances of gestural alignment and corresponding conceptual alignment, gestural and conceptual non-alignment, and less-clear cases that suggest a complex relationship between gesturing and conceptualization. The authors argue that interpreters' gestures may often reflect a blending of the interpreter's own viewpoint with that imagined to be held by the source speaker.

Monika Chwalczuk analyzed public service interpreting by investigating gesture production in a corpus of video recordings featuring 24 interpreters filmed in healthcare, educational, and police settings. While the interpreters' A language was either English or French, their B languages involved a range of 15 different ones. The gestural landscape (distribution of gesture functions used) and interpreters' gestural profiles (average distributions of different gesture functions) proved to be fairly similar across the settings, with pragmatic gestures being the most common in each case. Qualitative analysis revealed gestural alignment involving mainly deictic and representational gestures recruited in the processes of conceptual grounding, participatory sense-making, and disambiguation of lexical items. Through its consideration of patterns in gesture use, this research thus moves beyond previous descriptive, case-oriented studies on multimodal aspects of public service interpreting.

Inez Beukeleers, Laura Theys, Heidi Salaets, Cornelia Wermuth, Barbara Schouten, and Geert Brône studied the impact of gestural omissions and additions in interpreter-mediated medical encounters. The study involves data from interpreted consultations between two patients—one Russian-speaking and one Turkish-speaking—with Flemish Dutch-speaking healthcare professionals. A qualitative analysis of three excerpts from these interactions shows that omitting and/or adding representational and deictic gestures can potentially lead to changes in meaning, i.e., less or more concrete renditions of the original talk. The results highlight the semiotic complexity of healthcare interpreting, where visual elements (such as test results) and visualizable elements (including biological processes and medical procedures) are crucial topics of the discourse.

Vibeke Bø investigated embodied participation frameworks in one signed-to-spoken interpreted encounter between NTS and spoken Norwegian using embodied conversation analysis in a qualitative study. A crucial point here is that gaze is normally an important interactional resource in conversation, but in simultaneous signed-to-spoken interpreting, the interpreter's gaze is occupied with perceiving the signed discourse. Head gestures were therefore found to compensate as an alternative, and the research considers the various ways in which this plays out. The findings from this study highlight the need for further exploration of how interpreters navigate competing communicative demands.

Sílvia Gabarró-López also considered signed-to-spoken interpreting, but from LSFB into spoken Belgian French. The focus of the study is on the use of reformulation structures in the rendering of two LSFB dialogues and within the use of Belgian French as the target language. The most frequent forms of reformulation structures are found in both datasets, but a smaller number of reformulation structures was used in the target Belgian French discourse, likely due to factors such as the interpreters' cognitive load and the time pressure involved in producing

renderings (among other factors). In addition, while the interpreters drew on all their available semiotic resources to convey meaning, they did not seem to be influenced in their gesturing by the signs produced in the source LSFB dialogues.

It is our hope that this special issue will help draw attention to the value that gesture studies can bring to interpreting studies, and vice versa, the value that the domain of interpreting can have for future gesture research. In particular, we would highlight the potential for researching the vastly understudied discourse of those providing spoken interpretation of signed languages. Finally, we would like to thank the editors of *Parallèles* for being willing to consider this topic that takes interpreting studies in a new direction (the multimodal turn) and for their efficient communication process with us throughout the development of this special issue.

3. References

Anderson, L. (1979). *Simultaneous interpretation: Contextual and translation aspects*. Master's thesis, Concordia University.

Bø, V. (In press). Investigating interpreter-mediated interaction through the lens of depictions, descriptions, and indications: An ecological approach. *Translation and Interpreting Studies (TIS)*.

Bressem, J., Ladewig, S. H., & Müller, C. (2013). Linguistic annotation system for gestures. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body–language–communication: An international handbook on multimodality in human interaction* (Vol. 1, pp. 1098–1124). De Gruyter Mouton.

Cienki, A. (2013). Gesture, space, grammar, and cognition. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in language and linguistics: Geographical, interactional, and cognitive perspectives* (pp. 667–686). Walter de Gruyter.

Cienki, A. & Iriskhanova, O. K. (2020). Patterns of multimodal behavior under cognitive load: An analysis of simultaneous interpretation from L2 to L1. *Voprosy Kognitivnoy Lingvistiki, 1*, 5–11.

Davitti, E. (2019). Methodological explorations of interpreter-mediated interaction: Novel insights from multimodal analysis. *Qualitative Research, 19*(1), 7–29. https://doi.org/10.1177/1468794118761492

Gabarró-López, S. (2024). Towards a description of PALM-UP in bidirectional signed language interpreting. *Lingua 300*, 103646. https://doi.org/10.1016/j.lingua.2023.103646

Gile, D. (2009). Interpreting studies: A critical view from within. *MonTi: Monografías de Traducción e Interpretación, 1*, 135–155.

Janzen, T., & Shaffer, B. (2013). The interpreter's stance in intersubjective discourse. In L. Meurant, A. Sinte, M. Van Herreweghe, & M. Vermeerbergen (Eds.), *Sign language research, uses and practices: Crossing views on theoretical and applied sign language linguistics* (pp. 63–84). De Gruyter Mouton.

Kendon, A. (1972). Some relationships between body motion and speech. An analysis of an example. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communicatio*n (pp. 177–210). Pergamon Press.

Lang, R. (1976). Interpreters in local courts in Papua New Guinea. In W. M. O. Barr & J. F. O. Barr (Eds.), *Language and politics* (pp. 327–365). Mouton.

Lang, R. (1978). Behavioural aspects of liaison interpreters in Papua New Guinea: Some preliminary observations. In D. Gerver & W. H. Sinaiko (Eds.), *Language, interpretation and communication* (pp. 231 –244). Plenum Press.

McNeill, D. (1979). *The conceptual basis of language*. Erlbaum.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press.

Müller, C. (1998). *Redebegleitende Gesten: Kulturgeschichte—Theorie—Sprachvergleich.* Berlin Verlag Arno Spitz.

Nicodemus, B., Swabey, L., Leeson, L., Napier, J., Petitta, G., & Taylor M. M. (2017). A cross-linguistic analysis of fingerspelling production by sign language interpreters. *Sign Language Studies, 17*(1), 143–171. https://doi.org/10.1353/sls.2017.0000

Nilsson, A. L. (2015). Embodying metaphors: Signed language interpreters at work. *Cognitive Linguistics, 27*(1), 35–65. https://doi.org/10.1515/cog-2015-0029

Wang, J. (2021). *Simultaneous interpreting from a signed language into a spoken language: Quality, cognitive overload, and strategies*. Routledge.

Zagar Galvão, E. (2020). Gesture functions and gestural style in simultaneous interpreting. In H. Salaets & G. Brône (Eds.), *Linking up with video: Perspectives on interpreting practice and research* (pp. 151–179). John Benjamins Publishing Company.

Parallèles 37(1) April 2025

Making gestures inside and outside the booth: A comparative study

Celia Martín de León
Universidad de Las Palmas de Gran Canaria

Elena Zagar Galvão
Faculdade de Letras e Centro de Linguística da Universidade do Porto

Abstract

Although studies on multimodality in interpretation are gaining momentum, no research has been carried out comparing the multimodal behaviour of simultaneous interpreters inside and outside the booth to date. This exploratory study aims to compare the co-speech gestures and adaptors made by five professional conference interpreters while interpreting simultaneously and in face-to-face communication. The starting hypotheses are that participants will make more and larger gestures during the interview, and more adaptors during interpreting. Participants were filmed in both situations, and sections of similar duration of the videos were analysed and annotated with ELAN to obtain the gesture rate (number of gestures per minute), gesture amplitude, and adaptor rate (number of adaptors per minute) for each participant in each situation. The results obtained invalidate the first hypothesis (the gesture rate was higher in the booth in all cases), confirm the second hypothesis (the gestures were broader during the interview), and are inconclusive with respect to the third hypothesis. The analysis of the adaptors presented special methodological challenges that need to be further explored. The finding of a higher gesture rate during interpreting than during the interview might question the categorization of simultaneous interpretation as a monologic activity.

Keywords

Co-speech gesture, simultaneous interpreting, adaptor, face-to-face communication, gesture rate

 DOI:10.17462/para.2025.01.02

1. Introduction

Co-speech gestures, i.e., manual movements that are synchronous with utterances and are also related to them semantically and pragmatically (Gullberg, 1998; Kendon, 2004; McNeill, 1992) show great variability across individuals (Özer & Göksun, 2020). Moreover, language, culture, context, and communicative situations have been shown to influence their production (Alibali et al., 2001; Kita, 2009; Kita & Özyürek, 2003). Simultaneous interpreting (SI) is a complex form of cross-linguistically mediated communication and an extremely demanding cognitive and social activity, which has been aptly described as 'extreme bilingual language use' (Arbona et al., 2022, pp. 13). Simultaneous interpreters (hereafter simply 'interpreters') need to understand, analyse and process multimodal input in one language while at the same time producing and monitoring their verbal output in another language. While doing so, they display their own multimodal behaviour, which includes co-speech gestures (to a greater or lesser extent) as well as adaptors (Ekman & Friesen, 1969), i.e., hand gestures that seem to be independent of speech content, such as manipulating an object (other-adaptors) or touching a part of one's body (self-adaptors). Among the many questions that these preliminary considerations may prompt, two appear to be particularly relevant: (a) are interpreters' multimodal behaviours similar inside and outside the booth? and (b) to what extent is an interpreter's behaviour influenced by the multimodal behaviour of the speaker they are interpreting? This study seeks to provide potential answers to question a) through a quantitative analysis (gesture frequency) as well as a qualitative analysis (gesture amplitude) of interpreters' co-speech gestures and adaptors while they engage in SI and in face-to-face communication. The second question is currently being researched as part of an ongoing project but shall not be addressed here.

2. Gesture in simultaneous interpreting

Gesture in SI is an under-researched area in both Interpreting Studies and Gesture Studies. In Interpreting Studies, it is generally assumed that speakers' co-speech gestures are relevant for simultaneous interpreters' comprehension of the source speech, since they are an integral part of speakers' multimodal meaning-building and communication resources (Arbona et al., 2022; Galhano-Rodrigues, 2007; Poyatos, 1997; Sineiro de Saa, 2003; Zagar Galvão, 2015). A speaker's manual gestures and adaptors often supplement verbal content in various ways, thus helping interpreters confirm their comprehension of the source speech and even anticipate further speech content. In other words, interpreters need to see the speakers and, ideally, also the audience, to be able to glean all the elements that will help them deliver their professional service to the best of their abilities. Indeed, recommendations on visual input are routinely included in professional associations' guidelines on working conditions (e.g., AIIC Guidelines on working conditions). However, there seems to be little consensus about (or even awareness of) the role played by simultaneous interpreters' hand gestures. Some view these manual movements as a means to lighten interpreters' cognitive burden and facilitate their communication efforts (Cienki & Iriskhanova, 2020; Stachowiak-Szymczak, 2019), while others deem them undesirable, as they may distract from the main task at hand, which is eminently verbal and vocal (see Zagar Galvão, 2015). In Gesture Studies, simultaneous interpreting affords the opportunity to investigate multimodal behaviour in a unique communicative situation marked by the interaction of two languages, high cognitive load, as well as mental stress and time pressure.

Research on gesture in interpreting is very recent but has been gaining momentum, as attested by the panel on 'Gesture in spoken and signed-to-spoken language interpreting' at the last International Pragmatics Association conference (Brussels, July 2023). This one-day panel, convened by Sílvia Gabarró-López (Pompeu Fabra University and University of Namur)

and Alan Cienki (Vrije Universiteit Amsterdam), brought together scholars whose common denominator is the interdisciplinary study of gesture in interpreted discourse as a way to "shed light on how interpreters structure their discourse and on how their bodily actions construct meaning" (Gabarró-López & Cienki, 2023).

The earliest studies devoted to gesture and speech in interpreting appeared at the beginning of this century. Galhano-Rodrigues (2007) conducted a holistic micro-analysis of the relationship between speech, prosody, and co-speech gestures in a simultaneous interpretation from English into French using naturalistic data collected in a conference setting. Zagar Galvão adapted the methodology proposed by Galhano-Rodrigues and investigated the multimodal behaviour of trainee interpreters in a SI training session (2009) and by professional interpreters in a naturalistic and an experimental setting (2013; 2015). She also investigated interpreters' perception of their own gestural action while interpreting (2021).

The studies by Galhano-Rodrigues and Zagar Galvão showed that interpreters produce gestures to organise discourse, to highlight specific elements of discourse, to signal repairs and false starts, to represent semantically related verbal content, to express stance, and to regulate turns in the booth. Interpreters' co-speech gestures were shown to be intimately connected to prosody and to facilitate both speech comprehension and speech production. An intriguing finding was the presence of a degree of gestural and conceptual alignment between interpreters and speakers in all datasets, i.e., interpreters produced gestures whose form and meaning were very similar to those produced by the speakers (Zagar Galvão, 2013). An experimental study with interpreting students (Adam & Castro, 2013) concluded that beats (McNeill, 1992), i.e., gestures accompanying the rhythm of speech, were the most frequent gestures and also the only type that all the students made. Chaparro Inzunza (2017) investigated the relation between gesture and interpreting quality in SI and observed that quality decreased when student interpreters were prevented from gesturing. Stachowiak-Szymczak (2019) used eye tracking technology as well as gesture analysis and annotation tools to investigate gaze and gesture in both simultaneous and consecutive interpreting, viewed as embodied multimodal language tasks. One of her main conclusions was "that the level of congruence between the visual and auditory input affected the frequency of looking at the experimental screen, which indicates that visual and auditory input are integrated in language processing and in interpreting" (Stachowiak-Szymczak 2019, p. 137). Fernández Santana and Martín de León (2021) studied interpreters' embodied cognition by exploring the role played by iconic gestures and mental images in meaning-building during SI. The same authors investigated the role of referential and pragmatic gestures during an interpreting session with a professional interpreter and found that referential gestures supported the construction of meaning, while pragmatic gestures helped to manage the flow of the interpretation (Martín de León & Fernández Santana, 2021).

Cienki and Iriskhanova (2020) compared the gestures produced by novice and experienced interpreters in an experimental setting, where the participants were asked to interpret a TED Talk from English into Russian, their A language. To our knowledge, this is the first study of gesture in SI to have included an analysis of the interpreters' adaptors as well as their co-speech gestures. The results of the quantitative and qualitative analyses revealed that beats, i.e., hand gestures that accompany the rhythm of speech (McNeill, 1992), and adaptors were the most frequently used by inexperienced as well as experienced interpreters. According to the authors, these findings suggest that interpreters may be relieving part of their cognitive pressure through these hand movements.

Arbona et al. (2022) conducted two experiments with 24 professional interpreters and a control group of 24 professional translators, using eye-tracking technology. One of their main

goals was to establish whether interpreters attend to speakers' gestures and actually integrate the information thus obtained to facilitate language comprehension. They concluded that co-speech gestures that are semantically connected to utterances can help interpreters process the source speech:

> [All] this suggests that co-speech gestures are part and parcel of language comprehension in bilingual processing even in 'extreme bilingual language use', such as SI. [. . .] Overall, the results strengthen the case for SI to be considered a multimodal phenomenon (Galvão & Rodrigues, 2010; Seeber, 2017; Stachowiak-Szymczak, 2019) and to be studied, taught and practised as such (Arbona et al., 2022, p. 13).

These findings are supported by neurolinguistic research in SI using functional MRI, which has revealed that a brain area specialised in hand movement is indeed activated in SI (Ahrens et al., 2010). The authors of this study even suggest that interpreters may benefit from "hand activation" or "auxiliary motor action" to manage and control the speed of speech production and the potential overload of an area of the brain (the left superior temporal sulcus, STS) which helps speech perception (Ahrens et al., 2010, p. 245).

What becomes apparent from this brief literature review is the rich potential of an interdisciplinary study of interpreting and gesture. However, as research on gestures in interpreting is an emerging field and the number of studies is limited, we still do not know whether simultaneous interpreters display a specific multimodal behaviour in the booth as compared to their multimodal behaviour in other communicative settings. One way to address this question is by comparing the multimodal behaviour of simultaneous interpreters inside and outside the booth. Furthermore, this comparison may also shed light on the individual gestural styles of simultaneous interpreters.

3. Exploring interpreters' gestures inside and outside the booth

3.1. Aims and hypotheses

The aim of this study is to explore multimodal behaviour in simultaneous interpreting, comparing the co-speech gestures[1] and adaptors made by five professional interpreters while interpreting simultaneously in the booth and in face-to-face communication. Previous research indicates that speakers exhibit a higher gesture rate—a greater number of gestures per minute—and a larger gesture amplitude when addressing a visible interlocutor than when the interlocutor cannot be seen; and that gesture rate is also higher when an interlocutor exists, but is not visible (e.g., telephone conversations), than when there is no interlocutor at all (Bavelas et al., 2008). Furthermore, Cienki and Iriskhanova (2020) (cf. section 2 above) found some similarities across the gestures of simultaneous interpreters, such as the prevalence of adaptors and beats as opposed to representational gestures—gestures that depict some aspect of an utterance's content (Kendon, 2004, p. 160). All these considerations led to the formulation of the following initial hypotheses:

(1) the rate of co-speech gestures will be higher in face-to-face communication than during interpreting;

(2) co-speech gestures will be larger in face-to-face communication than during interpreting;

(3) the rate of adaptors will be higher during interpreting than in face-to-face communication.

[1] Simply put, co-speech gestures are hand movements where one or both hands depart from a resting position, achieve a salient configuration called 'stroke' and go back to a resting position where the movement ends.

3.2. Design of the study

To test the above hypotheses, a comparative analysis was conducted on the multimodal behaviour exhibited by five professional interpreters while interpreting simultaneously and while engaging in face-to-face dialogue. The research builds partially upon the work of Zagar Galvão (2015), who conducted an experimental study involving four professional interpreters in two distinct interpreting scenarios: one featuring a speaker with minimal expressiveness, and another with a highly expressive speaker (cf. Table 1 for a summary of each speaker's delivery style[2]). The present study draws on the data collected in the second scenario, which were reanalysed and expanded by adding data from a fifth participant. Additionally, the multimodal behaviour of all five participants during an interview session was examined.

Delivery style	Speaker 1	Speaker 2
Degree of expressiveness	Low	High
General body posture	Sitting down at a desk	Standing and moving on a stage
Gesture rate (gestures per minute)	~19.8	~ 33.4
Gesture quality	Predominance of pragmatic gestures produced mainly within the center and periphery of the speaker's gesture space	Higher number of both referential and pragmatic gestures produced within the center and periphery as well as in the extreme periphery of the speaker's gesture space
Prosodic features	Low degree of variation	High degree of variation

Table 1. Speakers' delivery style (Zagar Galvão, 2015, pp. 147–148)

Five professional conference interpreters (two women and three men) participated in the study. All have over twenty years' experience and are members of interpreters' associations. One of them is an EU-accredited freelancer, and another is a permanent staff member at the European Commission. Their A languages are Portuguese (4) and Italian (1). All share English as their B language.

As mentioned above, the multimodal behaviour of each participant was analysed in two distinct settings: while interpreting simultaneously and during an individual interview. In the first setting, each participant interpreted the video of a TED Talk by neuroscientist Vilayanur Ramachandran about the correlation between brain damage and cognitive functions[3]. The video is a recording of a real talk in which the neuroscientist addresses a fairly large audience. This 23-minute talk was interpreted from English into each participant's A language (i.e., from English into Portuguese by four interpreters and into Italian by one interpreter). Each participant interpreted alone in the booth without an audience. In the second setting, each participant was interviewed by one of the researchers following the same script.

3.3. Data collection

A remote interpreting assignment was simulated for data collection in the interpreting setting. Before participants began interpreting, they were given a list of five technical terms that would appear in the video (Capgras syndrome, parietal lobe, fusiform gyrus, phantom limb, amygdala). In order not to influence their behaviour by mentioning gestures, they were told that the aim of the study was to investigate remote interpreting. The video was displayed on

[2] The gesture rates indicated in Table 1 do not include adaptors.

[3] The video can be downloaded from the TED Talks page and used for educational and research purposes (https://archive.org/details/VilayanurRamachandran_2007).

a computer screen placed on the table inside the booth, very close to the interpreters. The sound quality was rated as very good by all the participants. During the interpreting session, participants were filmed using a small digital video camera located to their right inside the booth, except for participant G, who was filmed with a camera positioned outside, in front of the booth. Though a camera is an intrusive element for interpreters, most of whom do not like to be recorded on audio, let alone on video, at the end of the session, all the participants reported that they had completely forgotten about the camera filming them.

The five participants were then interviewed using a semi-structured script. The interviews were conducted in each participant's A language and unfolded as informal conversations between peers. This was made possible by the shared professional background between the interviewer, who is also a conference interpreter, and the interviewees. The conversations were recorded with a video camera that captured both the interviewer and the interviewee. Written consent was obtained from all parties involved for the recordings. Additionally, the five interpreters participating in the study signed written authorizations for the use of digital and printed images taken from the video recordings.

3.4. Data analysis

The initial seven minutes of the interpretation of the TED talk were transcribed[4], analysed, and annotated using ELAN 6.4 (Sloetjes & Wittenburg, 2008). Additionally, the interviews were transcribed, and a section (consistent across all participants) was selected for analysis and annotation. The duration of the interpreters' speaking turns in this section approximately matched the duration of the analysed portion of the interpretation (see Table 2).

	Booth	Interview	
Participant	time analysed	time analysed	speaking turns
A	07:10	09:00	7:00
G	07:08	10:45	7:00
I	07:13	11:00	7:42
J	07:08	09:00	7:00
M	07:06	08:07	7:27

Table 2. Time analysed in each setting in minutes

Since the aim of the research was to compare the multimodal behaviour of each interpreter in two different situations, the analysis was based on the concept of "interpreters' gestural style" proposed by Zagar Galvão (2015, 2020) to describe the interpreters' individual gestural profile. Building upon the notion of "interpreting style" by Van Besien and Meuleman (2008), Zagar Galvão introduced three continua that can be used as analytical tools to describe the general gestural style of an interpreter: gesture frequency (total number of gestures, rate of gestures per minute, or rate of gestures per 100 words), gesture size, and gestural mimicry, i.e., the conceptual and/or formal alignment of an interpreter with the speaker's gestures. Each of these continua describes one dimension of the interpreters' gestural behaviour. Depending on the specific objectives of each research project, new continua can be added to describe other relevant dimensions.

Following Bavelas et al. (2008), who suggest that speakers tend to use more gestures and broader gestures when addressing a visible interlocutor, this study focussed on the first two

[4] Fluent transcription (Setton 2002, p. 32).

continua. Additionally, a new dimension—the rate of adaptors—was introduced to explore whether adaptors are frequently used in simultaneous interpreting, as observed by Cienki and Iriskhanova (2020). Thus, the analysis was based on the following dimensions:

(1) gesture rate (number of gestures per minute, GPM);

(2) gesture amplitude (small, medium, large), measured according to their location and trajectory in gesture space (McNeill, 1992);

(3) adaptor rate (number of adaptors per minute, APM).

When calculating the gesture rate, multiple and repetitive strokes (i.e., a sequence of strokes with the same hand shape, orientation, trajectory, and performed within the same location in gesture space) were counted as one single stroke (Zagar Galvão, 2015, p. 153). Gesture amplitude was calculated following McNeill's division of gesture space into centre-centre, centre, periphery, and extreme periphery (McNeill, 1992, p. 89). Gestures were labelled as small when performed in only one area, medium when performed across two areas, and large when performed across three or more areas of the gesture space (see Figure 1). A gesture was classified as an adaptor when a participant manipulated an object or touched a part of his/her body.

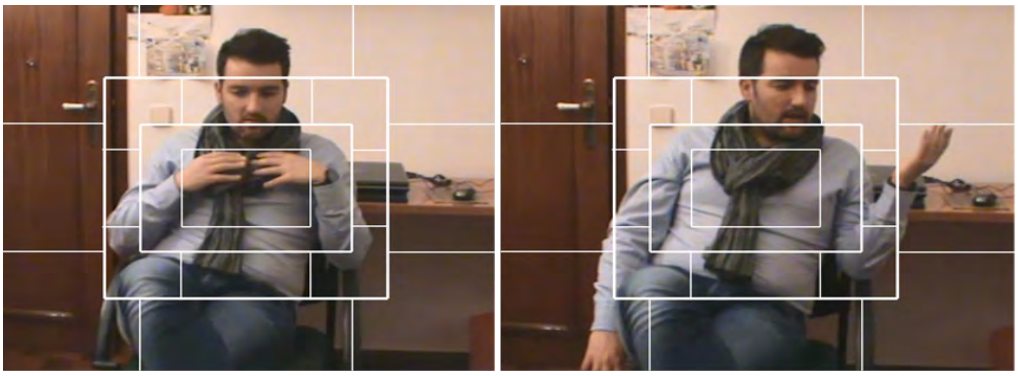


Figure 1. Division of gesture space following McNeill (1992)

The identification and annotation of gestures and adaptors was conducted separately and independently by each researcher. For each section, one researcher analysed the complete sequence, while the other analysed a fragment that accounted for between 20% and 100% of the entire sequence (see Table 3, 'time'). Several meetings were held before and during the analysis to align criteria and address doubts. At the conclusion of this process, the percentage of inter-rater agreement was calculated for both gesture and adaptor counts. The basis for calculating this percentage was the total number of gestures identified in each fragment by the researcher who analysed the entire sequence in each case ('N' in Table 3). As can be seen in Table 3, the percentage of inter-rater agreement exceeds 90% in all cases for gestures, but is lower in most instances for adaptors. This lower inter-rater agreement may suggest a greater difficulty in identifying adaptors compared to identifying gestures, an issue which will be addressed below. However, it should also be noted that inter-rater agreement for adaptors is still quite high, indeed higher than 80% in most cases.

Part.	Gestures interview			Gestures booth			Adaptors interview			Adaptors booth		
	time	N	%	time	N	%	time	N	%	time	N	%
A	07:19	112	94.11%	07:07	208	96.63%	06:27	24	100%	02:18	24	77.77%
G	02:17	37	97.36%	07:00	113	93.38%	03:37	23	85.18%	07:07	23	97.56%
I	06:19	58	98.3%	01:51	28	93.33%	03:03	20	86.95%	07:13	20	88.88%
J	01:37	32	94.11%	01:59	27	100%	05:25	19	86.36%	02:28	19	91.66%
M	03:26	64	96.96%	07:06	221	93.30%	02:29	22	81.81%	07:06	22	88.88%

Table 3. Inter-rater agreement

4. Results and discussion

The results of the study will now be analysed and discussed in relation to each of the three initial hypotheses.

4.1. First hypothesis

The first hypothesis was that the rate of co-speech gestures per minute would be higher in face-to-face communication than during interpreting. However, the results did not support this hypothesis. Surprisingly, the rate of gestures was actually higher in the booth setting for all participants (see Table 4 and Figure 2). The gesture rate of A and I in the booth was more than double their rate in the interview setting. The difference was minimal in the case of G, who also displayed the lowest gesture rate in the booth. M, on the other hand, had the highest gesture rate in both settings. Despite these individual differences, the findings suggest that interpreters do not always gesture less while interpreting than in face-to-face communication.

Participant	GPM in the booth	GPM in the interview
A	29.02	14.44
G	16.14	14.28
I	22.72	9,87
J	25.23	16.57
M	31.12	20
average	24.84	15.03
median	25.23	14.44
SD	5.85	3.69

Table 4. GPM in the booth and in the interview

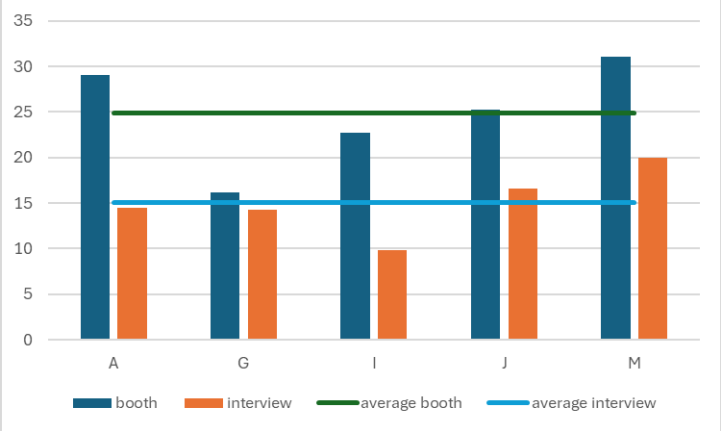


Figure 2. GPM in the booth and in the interview

How can this be explained, especially if one considers that simultaneous interpreting is a highly controlled form of discourse requiring constant input analysis and output monitoring? A possible reason for the higher gesture rate in the booth could be that cognitive load during simultaneous interpreting is higher, with gestures serving as cognitive support for the interpreter (Cienki & Iriskhanova, 2020; Ren & Wang, this special issue; Stachowiak-Szymczak, 2019). Another reason could be the experimental setting itself: while interpreting, participants looked at the screen without distractions, which could increase their concentration on the speaker; in this case, the speaker had a dynamic delivery style and made a high number of

gestures, which could have prompted the participants to do the same. In fact, in many cases, participants aligned their gestures with those of the speaker (Zagar Galvão, 2015).

4.2. Second hypothesis

The second hypothesis was that co-speech gestures would be larger in the interview than in the booth. The overall results support this hypothesis, though there are some exceptions. As illustrated in Table 5 and Figure 3, the percentages of small gestures were higher in the booth than during the interview, although the difference was nearly insignificant in the case of M. Conversely, the percentages of medium gestures were higher during the interview than in the booth, except for M, who exhibited a higher percentage of medium gestures in the booth (see Figure 4). Finally, the percentages of large gestures were higher during the interview for all participants, except for G, who displayed a slightly higher percentage in the booth (see Figure 5).

setting	participant	small %	medium %	large %
booth	A	63.78	32.97	3.24
	G	46.27	48.73	5.04
	I	67.91	26.86	5.22
	J	60	38.75	1.25
	M	41.5	47.5	11
	average	55.89	38.96	5.15
	median	60	38.75	5.04
	SD	11.43	9.36	3.64
interview	A	41.07	41.07	17.85
	G	23.28	73.97	4.1
	I	25.42	55.93	18.64
	J	33.69	50	16.3
	M	41.28	40.36	18.34
	average	32.94	52.26	15.04
	median	33.69	50	17.85
	SD	8.45	13.75	6.18

Table 5. Percentages of small, medium, and large gestures per participant and setting

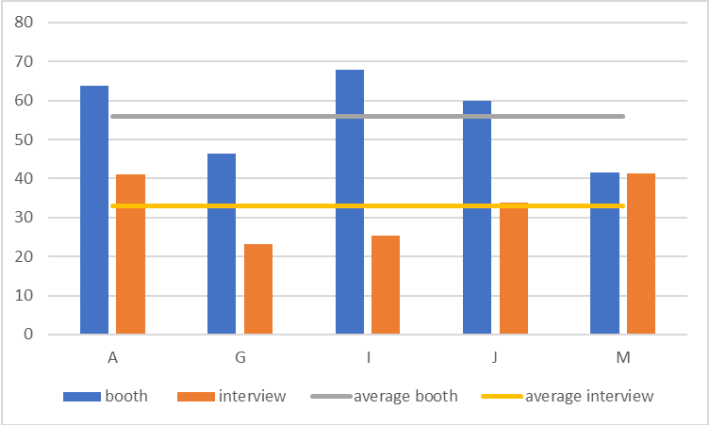


Figure 3. Percentages of small gestures per participant and setting

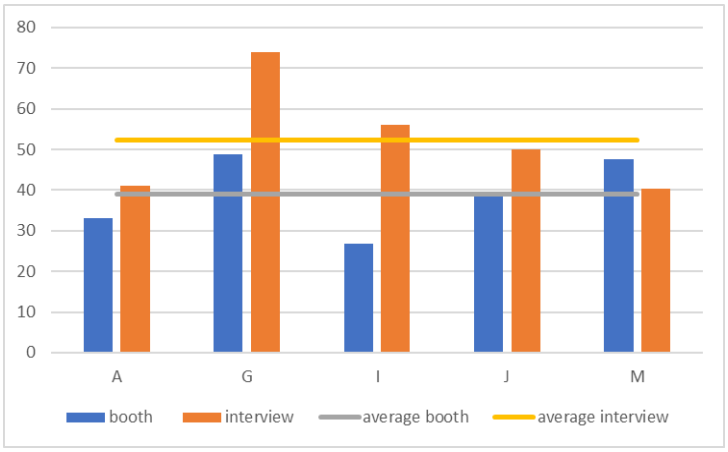


Figure 4. Percentages of medium gestures per participant and setting

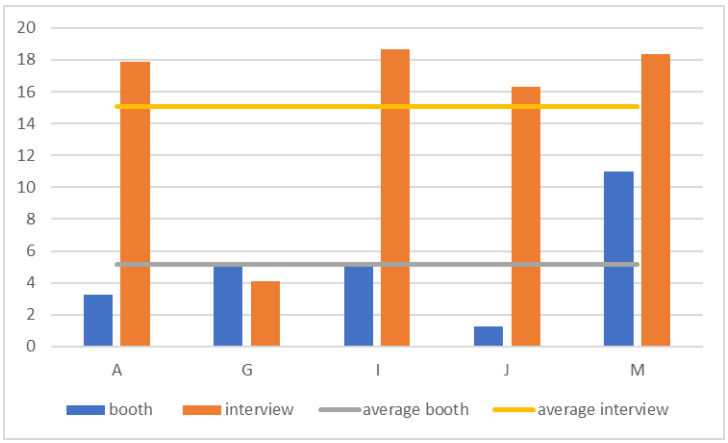


Figure 5. Percentages of large gestures per participant and setting

To better understand these differences, it is useful to compare the percentages of small, medium, and large gestures across both settings for each participant (see Table 5). Overall, participants exhibited greater gesture amplitudes during the interviews. However, Participant M showed smaller variations in gesture amplitude across the two settings, as illustrated in Figure 6. Thus, M's gestural profile is characterised by consistently high amplitude, exhibiting minimal variation across the two settings. As mentioned above, M also had the highest gesture rate in both settings. Both the high amplitude and the high gesture rate could indicate a gestural style marked by expressiveness and relatively low sensitivity to setting (see Figure 6).

Figure 6. Participant M gesturing in the booth and in the interview

Lastly, as mentioned above, participant G's multimodal behaviour deviates from the general trend observed, i.e., a higher percentage of large gestures in the interview setting as opposed to the booth setting. In his case, however, the percentage of medium gestures is much higher in the interview (see Figure 4). This discrepancy could potentially be attributed to the interview setup, as G's interview was the only one conducted at a table, a configuration which may have influenced his freedom of movement (see Figure 7).

Figure 7. Interview settings for M and G

Despite individual differences, the overall trend reveals a greater amplitude of gestures during interviews than during interpreting. This result aligns with Bavelas et al.'s findings (2008), where the presence of a visible interlocutor in the face-to-face dialogue situation resulted in larger gestures. In the current study, besides the presence of the interviewer and the dialogue situation, the space available for gesturing in each setting should also be taken into account. In the booth, interpreters' movements are constrained by the position they have to take at their 'workstation' (table, console, microphone) and the need to respect the space boundaries of their booth partner[5]. Conversely, during the interviews, the participants potentially had greater freedom of movement. Even in the case of participant G, whose interview took place in a classroom at a table, the available space for gesturing exceeded that of the booth (see Figure 8). Furthermore, during the interviews, a more relaxed atmosphere prevailed when compared to the tension often associated with interpreting. This laid-back ambience is reflected in the body posture of the participants, who often leaned against the back of their seats in the interviews, which may also have influenced the amplitude of their co-speech gestures.

[5] In this experiment, all participants worked alone in a booth. It is interesting to note, however, that they behaved as if they had a booth partner and strictly kept to their own workstation space.

Figure 8. Interpreting and interview settings for A, M and G

In summary, the data regarding the interpreters' gesture amplitude in the booth and during the interview support the second hypothesis and also highlight some striking individual differences. It should be noted, however, that gesture amplitude was measured from an external viewpoint without using any specific technology. Therefore, the results may not be as fine-grained as may be desired.

4.3. Third hypothesis

The third hypothesis, based on Cienki and Iriskhanova (2020), posited that the rate of adaptors per minute would be higher in the booth than during the interview. The results regarding this hypothesis are inconclusive. As can be observed in Table 6 and Figure 9, only participants A and G exhibited a significantly higher rate of adaptors per minute in the booth, while for participants I and M this rate was higher during the interview. It is noteworthy that the variability of the adaptor rate during the interview was low, with a standard deviation of 0.45. Among all the parameters analysed, this one exhibited the highest degree of uniformity across participants.

Participant	APM in the booth	APM in the interview
A	9.49	3.88
G	5.75	3.44
I	2.21	4.18
J	4.06	3.88
M	2.53	4.67
average	4.8	4.01
median	4.06	3.88
SD	2.97	0.45

Table 6. APM in the booth and in the interview

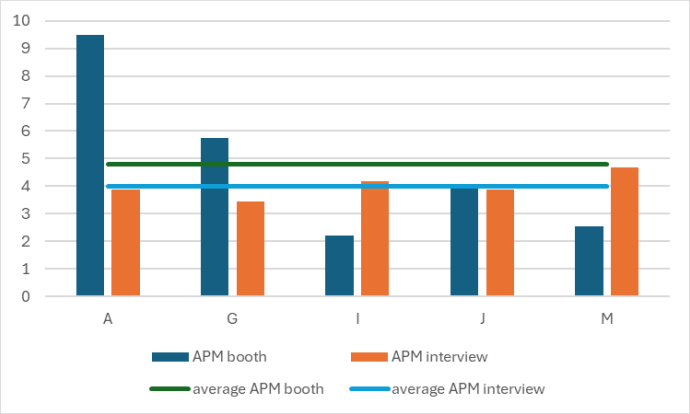


Figure 9. APM in the booth and in the interview

Nevertheless, the challenges encountered in annotating adaptors may have influenced the lack of conclusiveness of these results. Deciphering whether a movement constitutes a gesture or an adaptor, or distinguishing a resting position[6] from an adaptor, can be particularly daunting. Furthermore, segmenting a sequence of movements including gestures and adaptors into distinct units adds another layer of complexity. Little help could be obtained by consulting the literature, as there seem to be only a handful of studies that explicitly address the methodological difficulties associated with identifying and counting adaptors (e.g. Litvinenko et al., 2018; Żywiczyński et al., 2017). Moreover, adaptors are usually left out in studies about co-speech gestures and methodological guidelines, such as the Linguistic Annotation System for Gestures (Bressem et al., 2013, p. 1102). The following examples from the data are a good illustration of the difficulties associated with the analysis of adaptors.

In Figure 10, there is an example of a movement that could be categorised either as a co-speech gesture or as an adaptor. Interpreter A places his hand over his chest as he utters the bracketed part of the sentence "*Agora, [sou te sincero], há trabalhos em que é um alívio sair da cabine* (Now, [I'll be honest to you], there are jobs where it's a relief to leave the booth)." The synchronisation of the movement with the utterance led us to classify it as a gesture that conveys the meaning of sincerity (placing the open palm of your hand on your heart). However, in other contexts, a similar movement may be classified as an adaptor. For example, after the previous sequence, Interpreter A touches his scarf with both hands while saying "[*quer dizer*] *porque são chatos, porque não consegues fazer um trabalho bom* ([I mean,] because they [the speakers] are boring, because you [the interpreter] can't do a good job)" (cfr. Figure 11). In this instance, the movement was categorised as an adaptor because A appears to be adjusting the scarf, which he touches repeatedly, and there is no clear relationship between the movement and the utterance.

[6] Resting positions are the postures to which the hands return after performing a gesture.

Figure 10. Interpreter A's gesture in the interview: [*Sou-te sincero*] (I`ll be honest to you) (03:34.250 – 03:35.310)

Figure 11. Interpreter A's adaptor in the interview: *quer dizer* (I mean) (03:38.570 – 03:47.820)

Furthermore, adaptors can often be mistaken for the movement of one or both hands returning to the resting position after the stroke, as this motion often involves touching an object, a part of the body, or simply bringing one's hands together. Therefore, only those instances where repeated movement occurs were categorised as adaptors. For instance, in Figure 12, Participant A gestures with his right hand while saying "[*para mim*] ([for me])," and then returns to a resting position by holding a pen with both hands, saying "*este justificado… justificação não faz muito sentido* (this justified... justification doesn't make much sense)." In this specific case, Participant A holds and briefly spins a pen between the fingers of both hands before making the next gesture, and consequently the movement was categorised as an adaptor. However, such distinctions can be subtle, and in cases like this, it is often challenging to determine whether the movement is an adaptor or simply a return to the resting position.

Figure 12. Participant A's gesture and return to rest position in the booth:
[*para mim*] este *justificado… justificação* ([for me] this justified... justification)
(04:09.190 – 04:11.240)

Finally, when studying adaptors, it is challenging to divide movement sequences into distinct units, which clearly complicates quantitative analysis. For example, in Figure 13, Participant I performs an adaptor while listening to the interviewer. This adaptor lasts for 4.15 seconds, during which Participant I rubs and touches her hands in two different positions. First, she holds her right hand, palm vertical, fingers extended and slightly apart; then she wraps it around the thumb of her left hand. Despite its duration, this adaptor was considered a single one, since I's hands remained in her lap throughout. In instances like this one, where an adaptor has a fairly long duration, doubts arise about segmenting them into smaller units, especially when intermediate micro-pauses occur.

Figure 13. Participant I making an adaptor with two hand positions (03:05.990 – 03:10.140)

In the interview setting, adaptors often occurred while participants were listening to the interviewer. Both in the interview and during the interpretations, adaptors were mostly produced between gestures. Adaptors included other-adaptors, such as manipulating an object, and self-adaptors, such as touching a body part with one's hand/s. Sometimes, they were subtle movements preceding a co-speech gesture, serving as preparatory activations. Other times, they emerged during periods of doubt or hesitation. Some adaptors appeared purposeful, such as adjusting clothing or scratching a part of the body. Adaptors like these, whose purpose is unrelated to speech, have been termed "articulate," since they have a clear configuration and can be divided into phases. In contrast, adaptors lacking a specific purpose, called "subtle," lack articulation and internal structure and are considered signs of anxiety or stress, such as playing with a pen or rubbing one's fingers (Litvinenko et al., 2018, p. 7).

5. Conclusions

The invalidation of the first hypothesis suggests that simultaneous interpreting is a communicative situation in which, despite the absence of dialogue, the gestural rate may be higher than in dialogic situations. One possible explanation is that the cognitive load during simultaneous interpreting is higher and gestures serve as cognitive support for the interpreter. Cienki and Iriskhanova (2020) mentioned McNeill's curiosity when he observed an interpreter gesticulating in the booth while interpreting his lecture. In the absence of an interlocutor who could see the interpreter's gestures, McNeill inferred that "gesture was clearly being used for the speaker herself, not to communicate something to someone else" (Cienki & Iriskhanova, 2020, p. 7; see also Cienki, this special issue).

However, in addition to the cognitive pressure of the interpreting task, other factors may influence the gestural rate of interpreters, such as the expressiveness of the speakers and the extent to which the interpreters align with their gestures. Zagar Galvão (2015) identified a significant number of representational gestures made by the interpreters in her study, many of which exhibited common features with the speaker's gestures immediately preceding them, often used to describe objects and processes. This gestural alignment suggests that interpreters actively co-construct meaning with speakers and often adopt the speakers' point of view. In a sense, they become the speaker by 'owning the speech', a guideline that any interpreting student will have heard many times. These observations underscore the vital role of visual input as a variable in the study of interpreters' multimodal behaviour.

Finally, the experimental setting itself, with participants looking at a screen on which they could clearly see the speaker without any distractors, might have favoured their gesture production. In further research, it would be advisable to better simulate real working conditions in the experimental setting, for example by using remote interpreting tools. More research in naturalistic settings is also necessary.

The second hypothesis, stating that interpreters' co-speech gestures would be broader in the interview than in the booth, finds support in the overall results of the study. In general, the gestures of all interpreters were larger in the interviews, which can be attributed to the increased availability of free space for movement and a more relaxed atmosphere during these sessions. It is also conceivable that the face-to-face communicative setting and the presence of the interviewer prompted interpreters to employ more expansive co-speech gestures. In Bavelas et al. (2008), the face-to-face dialogue situation generated larger gestures, proportional to the size of the speaker's body. According to these authors (2008, p. 517), in narrative terms, this means that speakers in face-to-face dialogue situations adopt a "character viewpoint," with their hands representing the character's hands, and their body, the character's body (McNeill, 1992, p. 190). Although the present study does not specifically focus on a narrative context, comparing the narrative viewpoints represented in the participants' gestures across the two settings could be interesting.

However, the differences in the space available for gesturing in the booth and in the interview do not allow conclusions to be drawn about the influence that the presence of the interviewer and the dialogue situation in the second setting may have had on the amplitude of the gestures. In future research on gesture amplitude, it would be advisable to design the data collection so that both communicative situations occur in the same place or in similar places in terms of available space.

The results obtained in relation to the first two hypotheses also suggest that each interpreter has a particular gestural profile, as proposed by Zagar Galvão (2015). In particular, Participant M stands out for her high gesture rate and the amplitude of her gestures both in the booth and during the interview, exhibiting the least variation across the two settings.

Finally, the results pertaining to the third hypothesis (i.e., that the rate of adaptors per minute would be higher in the interpretation setting than in the interview setting) are inconclusive. Some participants produced more adaptors in the interview, others in the booth. Indeed, the challenges encountered when identifying and counting adaptors may well be the cause for this inconclusiveness. These challenges arise from the difficulty in distinguishing adaptors from co-speech gestures and return movements to resting positions, as well as in dividing adaptor movement sequences into units. Adaptors do not refer to the speech, and "subtle" adaptors are not structured. Unlike co-speech gestures, they do not have a stroke or culmination point. All this makes it more difficult to identify them precisely.

In addition to the methodological challenges encountered in the adaptor analysis, our study has other limitations. Firstly, the small number of participants restricts our ability to draw generalisations, thus the results should be viewed as exploratory and descriptive. Secondly, the calculation of gestural amplitude is not as precise as it would be if specific technology had been used.

Despite these hurdles, the study suggests two promising avenues for future research. First and foremost, there is a need for further exploration of adaptors and the methodological challenges associated with their analysis. This analysis can offer valuable insights into the mood and motivation of interpreters (Kendon, 2013, p. 9). Investigating the role of adaptors in SI and other interpreting modalities is a pending task, which is particularly urgent in consecutive interpreting, especially in high-stress situations, such as interviews at law enforcement agencies and immigration services.

One second promising avenue for further research relates to the gestural alignment of interpreters with speakers. Additional studies are needed to investigate whether interpreters' multimodal behaviour is communicative, and if their co-speech gestures, as part of their communicative activity, are aimed at their interlocutors, even if the latter do not generally perceive or attend to these movements. Some experimental studies on interpreters' gestures do not grant participants visual access to speakers (e.g., Cienki, this special issue; Cienki & Iriskhanova, 2020; Stachowiak-Szymczak, 2019). This lack of visual access to speakers could potentially influence the multimodal behaviour of interpreters, possibly resulting in a reduction in the number of representational gestures. It is crucial to consider this variable, especially now that a much greater number of SI jobs are carried out online through remote SI digital platforms. It is also important to continue investigating simultaneous interpreters' gestural alignment with speakers and its role in joint meaning-making, as well as researching the specific meaning of gestures in simultaneous interpreting. It may be concluded that the categorisation of simultaneous interpreting as a monologic activity has been challenged, a notion which certainly merits further research.

6. Funding information

Celia Martín de León's participation in this study was supported by a grant from the University of Las Palmas de Gran Canaria for a research stay at the University of Vienna. This grant was funded by the Spanish Ministry of Universities (Orden UNI/501/2021 of 26 May) and the European Union (Next Generation EU Funds).

7. References

Adam, C., & Castro, G. (2013). Schlaggesten beim Simultandolmetschen – Auftreten und Funktionen [Beat gestures in simultaneous interpreting – Occurrence and functions]. *Lebende Sprachen 58*(1), 71–82. https://doi.org/10.1515/les-2013-0004

Ahrens, B., Kalderon, E., Krick, C. M., & Reith, W. (2010). fMRI for exploring simultaneous interpreting. In D. Gile, G. Hansen, & N. K. Pokorn (Eds.), *Why translation studies matters* (pp. 237–247). John Benjamins. https://doi.org/10.1075/btl.88.20ahr

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188. https://doi.org/10.1006/jmla.2000.2752

Arbona, E., Seeber, K. G., & Gullberg, M. (2022). Semantically related gestures facilitate language comprehension during simultaneous interpreting. *Bilingualism: Language and Cognition, 26*(2), 425–439. https://doi.org/10.1017/S136672892200058X

Bavelas, J., Gerwing, J., Sutton, C. & Prevost, D. (2008). Gesturing on the telephone: independent effects of dialogue and visibility. *Journal of Memory and Language, 58*, 495–520. https://doi.org/10.1016/j.jml.2007.02.004

Bressem, J., Ladewig, S. H., & Müller, C. (2013). Linguistic annotation system for gestures. (LASG). In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body – language –communication. An international handbook on multimodality in human interaction*. *Vol. 1* (pp. 1098–1124). Walter de Gruyter. 10.1515/9783110261318.1098

Chaparro Inzunza, W. B. (2017). *Gesticulación y calidad de la interpretación simultánea. Un estudio experimental*. Master's thesis, Universidad de Concepción, Chile.

Cienki, A. Functions of gestures during disfluent and fluent speech in simultaneous interpreting. Parallèles, 37(1), 29-46. https://doi.org/10.17462/PARA.2025.01.03

Cienki, A., & Iriskhanova, O. K. (2020). Patterns of multimodal behavior under cognitive load: An analysis of simultaneous interpretation from L2 to L1. *Voprosy Kognitivna Lingvistika*, *1*, 5-11. 10.20916/1812-3228-2020-1-5-11

Eckman, F., & Friesen, W. V. (1969). The repertoire of non-verbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*(1), 49–98. https://doi.org/10.1515/semi.1969.1.1.49

Fernández Santana, A., & Martín de León, C. (2021). Con la voz y las manos: Gestos icónicos en interpretación simultánea – With voice and hands: Iconic gestures in simultaneous interpreting. *Hermēneus. Revista de Traducción e Interpretación*, *23*, 225–271. https://doi.org/10.24197/her.23.2021.225-271

Galhano-Rodrigues, I. (2007). Body in interpretation – Nonverbal communication of speaker and interpreter and its relation to words and prosody. In P. A. Schmitt & H. Jüngst (Eds.), *Translationsqualität* (pp. 739–753). Peter Lang.

Gabarró-López, S., & Cienki, A. (2023). Gesture in spoken and signed-to-spoken language interpreting. Panel description, *18th International Pragmatics Conference*, Brussels, 9–14 July 2023.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse. A study of learners of French and Swedish.* Lund University Press.

Kendon, A. (2004). *Gesture. Visible action as utterance.* Cambridge University Press. 10.1017/CBO9780511807572

Kendon, A. (2013). How the body relates to language and communication: Outlining the subject matter. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body – language –communication. An international handbook on multimodality in human interaction*. *Vol.1* (pp. 7–28). Walter de Gruyter.

Kimbara, I. (2006). On gestural mimicry. *Gesture*, *6*(1), 39–61. https://doi.org/10.1075/gest.6.1.03kim

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, *24*(2), 145–167. https://doi.org/10.1080/01690960802586188

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal: evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32. https://doi.org/10.1016/S0749-596X(02)00505-3

Litvinenko, A. O., Kibrik, A. A., Fedorova, O. V., & Nikolaeva, J. V. (2018). Annotating hand movements in multichannel discourse: Gestures, adaptors and manual postures. *The Russian Journal of Cognitive Science*, *5*(2), 4–17.

Martín de León, C., & Fernández Santana, A. (2021). Embodied cognition in the booth. Referential and pragmatic gestures in simultaneous interpreting. *Cognitive Linguistic Studies*, *8*(2), 277–306. https://doi.org/10.1075/cogls.00079.mar

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press. 10.2307/1576015

Poyatos, F. (1997). The reality of multichannel verbal-nonverbal communication in simultaneous and consecutive interpretation. In F. Poyatos (Ed.), *Nonverbal communication and translation* (pp. 249–282). John Benjamins. https://doi.org/10.1075/btl.17.21poy

Ren, Y., & Wang, J. Gesture and cognitive load in simultaneous interpreting: A pilot study. *Parallèles, 37*(1), 47-65. 10.17462/para.2025.01.04

Sineiro de Saa, M. (2003). Linguaxe corporal e interpretación simultánea: ¿Son os intérpretes un reflexo do orador? In L. Alonso Bacigalupe (Ed.), *Investigación experimental en interpretación de linguas: Primeiros passos* (pp. 39–59). Universidade de Vigo.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category—ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008) Marrakech, Morocco* (pp. 816–820). European Language Resources Association (ELRA).

Stachowiak-Szymczak, K. (2019). *Eye movements and gestures in simultaneous and consecutive interpreting*. Springer.

Van Besien, F., & Meuleman, C. (2008). Style differences among simultaneous interpreters. *The Translator, 14*(1), 135–152. https://doi.org/10.1080/13556509.2008.10799252

Zagar Galvão, E. (2009). Speech and gesture in the booth—A descriptive approach to multimodality in simultaneous interpreting. In D. de Crom (Ed.), *Translation and the (trans)formation of identities: Selected papers of the CETRA research seminar in Translation Studies 2008.* https://www.arts.kuleuven.be/cetra/papers (accessed 03 11 2023)

Zagar Galvão, E. (2013). Hand gestures and speech production in the booth: Do simultaneous interpreters imitate the speaker? In C. Carapinha & I. A. Santos (Eds.), *Estudos de linguística,* Vol. II (pp. 115–129). Imprensa da Universidade de Coimbra. 10.14195/978-989-26-0714-6_7

Zagar Galvão, E. (2015). *Gesture in simultaneous interpreting from English into European Portuguese: An exploratory study.* PhD dissertation, University of Porto.

Zagar Galvão, E. (2021). Making sense of gestures in simultaneous interpreting: The insiders' view. In C. Martín de León & G. Marcelo Wirnitzer (Eds.), *En más de un sentido: Multimodalidad y construcción de significados en traducción e interpretación* (pp. 118–143). Servicio de Publicaciones y Difusión Científica de la Universidad de Las Palmas de Gran Canaria. https://doi.org/10.20420/1653.2021.480

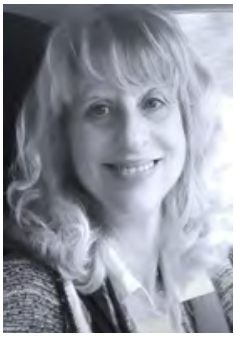
Zagar Galvão, E., & Galhano-Rodrigues, I. (2010) The importance of listening with one's eyes: A case study of multimodality in simultaneous interpreting. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds), *New insights into audiovisual translation and media accessibility* (pp. 239–253). Rodopi. https://doi.org/https://doi.org/10.1163/9789042031814_018

Żywiczyński, P., Wacewicz, S., & Orzechowski, S. (2017). Adaptors and the turn-taking mechanism: The distribution of adaptors relative to turn borders in dyadic conversation. *Interaction Studies*, *18*(2), 276–298. https://doi.org/10.1075/is.18.2.07zyw

Celia Martín de León
Universidad de Las Palmas de Gran Canaria
c/ Pérez del Toro 1
35003 Las Palmas de Gran Canaria
Spain

celia.martin@ulpgc.es

Biography: Celia Martín de León is an Associate Professor of Translation at the University of Las Palmas de Gran Canaria, Spain, where she has been teaching since 1995. Her research focuses on the empirical study of translation and interpreting processes from a multimodal, embodied, situated, and distributed perspective of cognition. She has explored meaning-making processes in translation and interpreting by addressing different phenomena such as conceptual metaphors, implicit theories, mental imagery, mental simulations, and co-speech gestures. Currently, her research is centred on the co-speech gestures of simultaneous interpreters. She has authored several articles and book chapters on these topics.



Elena Zagar Galvão
Faculdade de Letras da Universidade do Porto
Centro de Linguística da Universidade do Porto
Via Panorâmica s/n.
4150-564/Porto
Portugal

egalvao@letras.up.pt

Biography: Elena Zagar Galvão is an Assistant Professor of Translation and Interpreting at the University of Porto, Portugal, where she coordinates a Specialisation Course in Conference Interpreting and a Master's Degree in Translation and Language Services. Her research interests include multimodality, especially the relationship between speech and gesture in simultaneous interpreting, translator and interpreter training, and audiovisual translation. She is a member of the International Association of Conference Interpreters (AIIC).

Parallèles 37(1) April 2025

Functions of gestures during disfluent and fluent speech in simultaneous interpreting

Alan Cienki

Vrije Universiteit Amsterdam

Abstract

This study investigates what types (functions) of gestures occur during disfluencies in speech production during simultaneous interpreting as compared with gesture use during fluent interpreting. Forty-nine participants interpreted two ten-minute audio segments of popular science lectures, one from their first language to their second language and one from their L2 to their L1. The results show that during both fluent and disfluent moments of interpreting, the participants primarily used pragmatic gestures (such as marking emphasis) and self-adapters (e.g., rubbing their fingers). We can conclude that this points to the potentially different kind of thinking that is involved in speaking for simultaneous interpreting than is normally involved in thinking for spontaneous conversation or unrehearsed narratives. Self-adapters may assist the interpreters in the presentation of ideas and help with speech production. The low use of representational gestures may reflect the lack of deep semantic processing during simultaneous interpreting—not the kind of rich mental simulation which might give rise to depiction in gesture—and be a factor of the temporal constraints that do not allow for producing detailed gestural forms. Future research could involve comparison of gestures used by interpreters accompanying their own spontaneous speech with those they use while interpreting.

Keywords

Simultaneous interpreting, disfluencies, pragmatic gestures, self-adapters, thinking for speaking

 DOI:10.17462/para.2025.01.03

1. Background

The realization of the role of gesture in relation to spoken-language interpreting can be traced back to as early as 1971. It was then that the cognitive psychologist and psycholinguist, David McNeill, while giving a talk at a conference in Paris, took notice of a simultaneous interpreter, working in a soundproof booth in the back of the room, who was interpreting his lecture from English into French. He later wrote (McNeill, 2005, p. xi), "I could see a young woman behind the glass vigorously moving her arms in an alarming way," and his realization that this was because she was interpreting, or at least because she was speaking, helped determine the focus of his future research. "I believe I saw then, in a sudden apprehension via this distant yet strangely intimate connection of my speech to another person's movements, that language and gesture were two sides of one 'thing'" (p. xi). This interest led to McNeill developing a lab for gesture research at the University of Chicago whose ground-breaking work helped give rise to the modern field of gesture studies.

One of McNeill's (1992) seminal claims is that during spontaneous talk, our ideas develop and become "unpacked" not only in the words and grammatical forms that we speak, but also in bodily movements—gestures—of various kinds. McNeill named idea units "growth points", building on Vygotsky's (1934/1962) work that explained how, in the process of speaking, we are continually laying out new ideas against the background of ideas already known (either by having been uttered earlier or from context). As each idea unit arises, it is unfurled in speech and gesture, with information being verbalized in speech that can be fit into the linear lexico-grammatical system of the language that one is using, and with wholistic, imagistic information potentially appearing in the speaker's gestures. The production of speech and gesture works in a dialectical relation between the two forms of expression, with each potentially having an influence on the other. This is what is called the Growth-Point Hypothesis (McNeill, 1992; McNeill & Duncan, 2000). Kita et al. (2017) took this research a step further, arguing (based on their empirical studies) that it may not be the act of speaking per se that motivates the use of gesture, but more fundamentally, the formulation of concepts, particularly ones connected to spatial imagery. Since speaking one's own thoughts is based on such processes of conceptualization, gesture use is closely tied to what Slobin (1987) called "thinking for speaking".

In addition, there is a long tradition of research on gesture that concerns the communicative role of gestures, and it is in this tradition that Kendon (2016, p. 44) refers to gesture as "utterance dedicated visible bodily action". With this characterization he is building on several key points in his approach to studying gesture. First, it takes "utterance" as the starting point, viewing speaking a language, gesturing, signing a sign language, and potentially other actions as components of what one is doing when one is attempting to communicate. His choice of the term "action" distinguishes willful behaviors from uncontrolled ones (like spasms). "Visible" can be taken as meaning: available for perception as movement in space. The complement to this is then "utterance dedicated audible bodily action," which is how one could characterize speech. In this way, Kendon (1980) describes gesture (in his sense of "gesticulation") and speech in the sub-title of that paper as "two aspects of the process of utterance."

1.1. On gesture in interpreting

One of the unique aspects of interpreting is that the idea units that are being rendered do not stem from the interpreters themselves, but rather they come from someone else. In a sense, the idea units have to be reconstituted in the interpreter. This process is clearly different from that of how ideas for discussion come to one's mind when engaged in spontaneous conversation with someone else. Another particularity is that those that hear a spoken language interpreter's

audible action (the spoken renderings) may not see the interpreter's visible actions. This is particularly the case when the interpreter is located in a booth in the back of the room, behind the listening audience, which is a common arrangement for conference interpreting.

Consequently, spoken-language interpreting presents a unique context for the study of gesture use, in light of both the cognitively different motivation for interpreters' speech from that of speakers' self-generated talk and the interactionally different placement of interpreters, often working in a booth and out of view of those to whom they are speaking. In addition to this, while the mental effort exerted in speaking in a conversation is regulated by those engaged in the interaction (that is, a listener can facilitate the speaker's utterance production via co-construction or by providing feedback), the cognitive load of the simultaneous interpreter is known to be particularly heavy (Gile, 1997; Seeber, 2013). Beyond the fact of engaging in listening to new information while uttering information that had just been heard, the time constraints on keeping up with the interpreting, dependent on the rate of speech having to be interpreted (among other factors), is an additional demand on the task.

Previous research has begun to address a few of the ways in which gesture use relates to the process of simultaneous interpreting (SI). While the work to date provides fascinating insights, the findings so far have been limited in terms of the number of interpreters studied and disparate in their foci. Galhano-Rodrigues was one of the initiators of research in this field. In her 2007 study involving close description and analysis of the work of one simultaneous interpreter, she pointed out the important role that beat gestures played in the process of interpreting, movements which normally serve the pragmatic function of indicating emphasis. Though the interpreter in question here produced gestures serving different functions, beats, being aligned with prosodic stress, appeared to serve as a kind of "motor impulse" (p. 750), helping drive the interpreting process. However, since it was a qualitative study of one individual's performance, it is not possible to draw conclusions about interpreters' gestural behavior more generally. Zagar Galvão and Galhano-Rodrigues (2010) investigated two minutes of a session by one interpreter viewing the video of the speaker whom he was interpreting, considering whether he would imitate the gestures of the person speaking the source text. They found imitation of some of the original speaker's gestures on a small scale, but also some of the speaker's emphasis expressed in manual gestures was reproduced by the interpreter in other ways, such as with prosodic stress or head movements. Zagar Galvão (2015, 2020) researched two and four interpreters, respectively, with a similar goal and found that gestural imitation varied widely between the individual interpreters. This was a factor of individual differences between the interpreters in terms of both the quantity of gestures used and their functions (e.g., referential versus pragmatic functions). Martín de León and Fernández Santana (2021) examined gesture use when an interpreter looked at, versus looked away from, the video of the speaker being interpreted. Representational and deictic gestures appeared to support the construction and organization of meaning, while pragmatic gestures appeared to help manage the progress of the interpreting process. However, as an exploratory descriptive study, the research involved only one participant.

Overall, most of the research in this area to date has only considered very small numbers of participants for more qualitatively-oriented analysis. In addition, as initial explorations in this field, previous studies have had diverse goals, making it as yet difficult to draw broader conclusions. The following section lays out the motivations for the present study, which will focus on the question of gestures' potential role in relation to disfluencies during SI.

1.2. Motivations for the present study

1.2.1. On disfluency in interpreting

The heavy cognitive load that simultaneous interpreters experience in doing their work is known to result in various forms of disfluency in speech as they are rendering utterances in the target language (see Ren & Wang, this special issue). These include the truncation and restarting of utterances (Dayter, 2020; Gósy, 2007), the use of fillers like *uh(m)* (Plevoets & Defrancq, 2018), and long silent pauses (Ahrens, 2007). Particular elements in a source text/ speech being interpreted that are known to be more likely to lead to moments of disfluency include mention of numbers (e.g. Kajzer-Wietrzny et al., 2024), proper names, and the overall lexical density of the source text (Plevoets & Defrancq, 2016). Numbers, for example, are frequently interpreted incorrectly or are omitted (Mazza, 2001; Pellatt, 2006) in SI. This is due to factors such as their low predictability, the low redundancy in the information they convey, and yet the high information load constituted by them (Mazza, 2001; Pinochi, 2010). But as Plevoets and Defrancq (2016) point out, the cognitive demands in interpreting come not only from the source text (the "input load"), but also from the constraints on expressibility imposed by the target language (the "output load"), such as the grammatical forms available in it and what fixed phrases are frequently used in the language.

1.2.2. On the potential role of gesture in relation to disfluency in interpreting

There are several reasons to hypothesize that gestures serving different functions might play a role as interpreters attempt to resolve moments of disfluency in their speech. Here we will consider representational gestures, deictic gestures, gestures serving pragmatic functions, and self-adapters, as explained below. With the term "representational gestures" we are referring to use of one or both hands employing one or more of Müller's (1998a, 1998b, 2014) modes of representation. These involve either tracing a form, embodying a form, or acting as if touching or manipulating a referent that is mentioned in the co-gesture speech or that is inferable from the context of the talk. Interpreters might resort to representational gestures when trying to express concepts they have heard in the source text to help them with formulating that idea in the target language; witness the known role of depictive, iconic gestures in lexical retrieval (e.g., Krauss et al., 2000) and in information packaging that might aid in the lexicalization of concepts (Kita, 2000). Deictic (pointing) gestures are known to be used by speakers to identify referents in narration as they may point to different spaces to stand for different topics, referents, or times—what is known as abstract deixis (McNeill et al., 1993). This function of gesture could, in theory, aid interpreters in keeping track of ideas that they mentioned previously, or in differentiating new ideas by pointing to different spaces, thereby easing their cognitive load during disfluencies by offloading (Risko & Gilbert, 2016) some of the information onto the gesture space. Considering pragmatic gestures, one of the functions they are known to serve is that of word search—and of displaying in interaction the fact that one is engaged in searching for a word, thus helping the speaker hold the floor during an extended pause (Dressel, 2020). Some commonly recurring forms for such gestures are an open hand rotated at the wrist—the so-called cyclic gesture (Ladewig, 2011)—and a palm up (or diagonal) open hand (Clift, 2020; Müller, 2004). Finally, self-adapters[1] are self-touching movements such as rubbing one's fingers together, stroking one's hair, scratching oneself, etc. In particular, self-adapters involving sustained movement (e.g., a rubbing motion versus a simple one-time scratching movement) are known to help with maintaining one's mental focus and controlling stress (e.g., Ekman & Friesen, 1969).

[1] The American English spelling "self-adapter" is used here, but the British spelling "self-adaptor" is also common in the literature.

As the explanations above suggest, there is ample reason to suppose that any of these functions of gestures could aid simultaneous interpreters during moments of disfluency. The previous studies, discussed in section 1.1, do not yet provide a clear answer about this. A pilot study involving ten simultaneous interpreters (Cienki & Iriskhanova, 2020) did show self-adapters being used more than other gesture types, regardless of the fluency of the interpreting, but the distribution of other gesture functions was uneven across the participants, showing great individual variation. This leads to the research question for this study: What functions of gestures are used during moments of disfluency in SI and with what relative frequencies, and how does this compare to the functions and frequencies of gestures used during fluent SI? The answer to this question will contribute to the growing field of research on interpreting from a multimodal perspective.

2. Data collection

2.1. Participants

Two pools of participants were involved in the study. The first subset of data was collected in 2019-20 and involved interpreters working between Russian and English (N=29, 13 female), in both directions with different source texts. The second subset was collected in 2020-21 and entailed interpreting between Russian and German (N=20, 7 female), also in both directions. English and German were chosen as two languages which are from the same Indoeuropean language family (Germanic) but which have syntactic differences in the structuring of verb phrases, thus potentially providing a greater variety of reasons for disfluencies to arise in interpreting to and from Russian, a language relying more on pragmatic motivations for word order. All participants were native speakers of Russian and were either in training or working as professional simultaneous interpreters. Though each group consisted of interpreters with a range of experience, the results obtained in this study did not differ depending on the amount of experience, after we compared the results of those with three or more years of interpreting experience to those with less than three years' experience. Therefore, this factor was not taken into account any further in the study.

2.2. Stimuli for data collection

All participants interpreted excerpts from educational lectures about biodiversity and the extinction of species that were presented for the general public (laypeople) (see details in the section on the Procedure below). Those interpreting between Russian and English heard part of a lecture in Russian from the popular science website Postnauka entitled "Is there a threat today of a sixth mass extinction of species?"[2] which they interpreted into English, and also part of a TED Talk in English on "Mass extinction and the future of life on Earth"[3] which they interpreted into Russian. Those interpreting between Russian and German heard the same part of the same lecture in Russian noted above but interpreted it into German, and they also heard a portion of a lecture in German from the ARD TV website on "The end of evolution"[4] which they interpreted into Russian.

[2] "Существует ли сегодня угроза шестого массового вымирания видов?" https://postnauka.ru/video/49851, lecturer: Nikolai Dronin.

[3] https://www.ted.com/talks/michael_benton_mass_extinctions_and_the_future_of_life_ on_ earth?language=en, lecturer: Michael Benton.

[4] "Das Ende der Evolution" https://www.ardmediathek.de/video/tele-akademie/prof-dr-matthias-glaubrecht-das- ende-der-evolution/swr/Y3JpZDovL3N3ci5kZS9hZXgvbzEyMDkzOTk/, lecturer: Matthias Glaubrecht.

2.3. Procedure

Several days before coming in for their interpreting session, participants were provided with two glossaries, one per video, of about 20 discipline-specific terms that were used in the lectures, with possible translation equivalents for each term in the relevant target language. After obtaining informed consent from participants to take part in the study, they were only told that we were interested in the process of interpreting; they were informed of our interest in gesture research in a debriefing after their interpreting was completed. Each participant was brought to a booth used for training interpreters at Moscow State Linguistic University. They were not allowed to bring any materials with them, such as the glossaries, any paper or pens, or their phones. While this does not completely replicate interpreters' authentic conditions, we implemented this constraint so as to research how interpreters would handle the cognitive load of their task using only what Gibbon (2005) calls one's natural media—one's own body as a resource. In the booth they listened with headphones to the audio recordings to be interpreted, which were played on a laptop out of the interpreter's view. It is important to note that the participants only *heard* the portions of the lectures; they were not shown any video of the speakers. Before each turn at interpreting (Russian to English/German or English/German to Russian), they first heard a one-minute excerpt from the lecture so that we could properly adjust the volume to their choosing and so that they could become accustomed to the speaker. After that, they heard and interpreted the ten minutes of the lecture that followed the sample segment. The order in which the interpreting was performed (to or from Russian) was counterbalanced, differing randomly per participant. During the interpreting, they were left in the booth and the researcher sat in a nearby booth so that they could not be seen. The interpreter therefore looked out of the glass door of the booth into an empty classroom. After completing the two interpreting tasks, the participants filled in a second consent form, specifying how their video-recorded image could be shown in publications (choosing whether only as anonymized drawings or as screen shots/photos) and whether or not video clips could be shown at academic conferences or posted on academic websites in connection with publications of the research results.

2.4. Recording set-up

Each interpreter sat on a chair in front of the small desk in the interpreting booth. Participants were recorded from three angles. A Sony videocamera (recording at 25 fps) was placed on a tripod and positioned behind the seated interpreter, to the right side, such that the view it provided looked over the interpreter's right shoulder onto the desk, where the interpreters' arms and hands were. This afforded clearly seeing the forward and lateral movement of the interpreter's hands. In addition, a small GoPro camera (25 fps) was placed on the far edge of the desk in front of the interpreter, facing them. This recorded a close-up view of the interpreters' hands and also their face.

3. Methods of analysis

The videos and audio from the three cameras were synchronized and combined into one composite video for each interpreting session. Each composite video was imported into the software ELAN[5] (Sloetjes & Wittenburg, 2008) for analysis. This involved transcription of the speech and coding it for disfluencies, annotation of the gestures, and coding of them for their functions, as described below. Given the large amount of data obtained from the two interpreting sessions of each of the 49 participants, we selected two minutes from each

[5] https://archive.mpi.nl/tla/elan, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands.

session for coding and analysis, providing us with 196 minutes of analyzed data. The two specific parts chosen for analysis from each ten-minute session were minutes 3:00-3:59 and 8:00-8:59, taken as samplings of different points in the task. These were chosen as random portions after the interpreter had gotten into the interpreting task (not the initial minutes) and yet before the very end when the interpreter might have been more fatigued. Nevertheless, all the interpreters were hearing the same minutes of the respective lectures (the portions spoken in minutes 3 and 8) and thus within each language, they heard the same content.

The ELAN files were randomly distributed among three teams involved in the analysis for the project, with each team comprised of three coders. In each team, the annotation and coding described below was performed independently by each of the three team members, followed by a consensus check within that team. In addition, second coding was performed by one of the other teams, randomly assigned, for the presence of and the type of disfluencies in speech and the presence of and the functions of gesture phrases. Disagreements in coding were discussed and resolved at regular research group meetings, resulting in cross-checked files which were used to obtain the results.

3.1. Analysis of speech

The interpreters' renderings were annotated for moments of disfluency, coding for the following categories:

- Truncation. This involved suddenly cutting off a word or phrase, including when an utterance was begun, but abandoned (including "false starts") (Du Bois et al., 1993).
- Restart. This involves beginning an utterance again after a truncation (Du Bois et al., 1993). Some utterances were restarted more than once, in which the non-final attempts were also truncated. These were simply counted as restarts.
- "Stumbling". This was our collective term for instances of stuttering or mumbling. Stuttering involved quickly repeating a sound in a word in an apparently uncontrolled way. Any rapidly repeated truncations were counted here. Mumbling involved speaking for a short time in a low, indistinct manner or quickly saying a series of pieces of words.
- Filler. This was the term we used to cover words such as *well* in English or *nu* in Russian and non-lexical sounds (like *uh*, *uhm*), which Du Bois et al. (1993) call "marginal words".
- Dragging out of words or sounds or markedly slower tempo of speech, given the interpreter's rate of speaking otherwise.
- Long pauses were not determined based on absolute time criteria, as their length can vary per interpreter. They were only annotated as such if they were immediately followed by a stretch of very fast speech (catching up) or clear omission in the interpreting of the source text. Otherwise, pauses were not counted, as they constitute part of the normal process of uptake of information from the source text (Ahrens, 2007).

The remaining, non-disfluent interpreting is what we called fluent interpreting. It should also be mentioned that any time the interpreter replicated disfluencies on the part of the speaker of the source text (the original lecturer being interpreted), this was not coded, but such instances were also extremely rare.

3.2. Analysis of gestures

Given the particular role of manual gestures in relation to speech known from the literature in gesture studies and with the aim to delimit the scope of the study for practical reasons, only gestures of the forelimbs (hands and arms) were studied. The unit of analysis chosen was the gesture phrase (Kendon, 2004, ch. 7). This consists of the gesture stroke and any hold that

occurs after it. The stroke is defined here as a dynamic phase of clear, effortful movement, usually with an apex of movement. A post-stroke hold occurs when the hand "is held still in the position it reached at the end of the stroke" (Kendon, 1980, p. 213).

Gesture phrases were then coded for one of the several possible functions noted below. Although gestures are often multifunctional (Kok et al., 2015), we focused on our assessment of the most prominent or primary function of each gesture phrase. The relation of the gesture to the speech was taken into account in determining the gesture function. If the gesture phrase involved two hands, and the hands were not working in a complementary fashion (creating a two-handed gesture), the gesture of the speaker's dominant hand was coded, that being for our purposes the hand with which the speaker gestured the most during the interpreting session. We employed the following categories for determining each gesture's primary function.

- Representation involves depiction of some form or action. This was determined using an adaptation of Müller's (1998a, 1998b) "modes of representation". That is, if a gesture phrase primarily appeared to fulfill, given the speech and context in which it occurred, one of the modes of representation described by Müller, it was coded as representational. The modes were just used as a means of making a decision about representation or not; we did not perform analyses in relation to the individual modes. We used the following categories. Acting encompasses moving in a way in which the hand would normally perform some function, such as when a clasped hand is rotated as if turning an object around. Molding involves moving as if touching the surface of something, thereby showing its shape. Holding entails one or both open hands, usually with palm and fingers slightly cupped, briefly sustaining a position in space, as if holding something. In Tracing, one or more extended fingers move to show the outline of something with the fingertips. In Embodying, the hand takes on the form of the thing represented, involving displaying the hand or fingers in the shape of the referent, as when one's extended index and middle fingers alternately move back and forth to represent a person walking.
- Deixis involves specifying a referent or a spatial or temporal location from the perspective of either the situation described or the surrounding discourse. This can be accomplished by pointing with extended fingers or by touching (e.g., the interpreter tapping the desk in front of them with one or more extended fingers).
- Pragmatic gestures: for our purposes, this category was reserved for gestures which were not seen as primarily involving Representation or Deixis. As Kendon (2004, p. 158) notes, pragmatic gestures relate to features of what the speaker is expressing that are not part of the referential meaning of the utterance. This encompasses showing one's stance towards a topic (e.g., by shrugging), making emphasis (with a beat), indicating negation (with a lateral sweeping movement of the open hand), etc. This category usually involves gestures that recur across speakers and contexts within a given culture with similar function ("recurrent gestures" as described in Bressem and Müller, 2014; see also Grishina, 2017, ch. 14, on the pragmatic gestures frequently used in Russian culture).
- Adapters, for our analysis, can be self-adapters or other-adapters. Self-adapters involve a form of self-touching. This can include scratching oneself, rubbing one's own fingers or hands, or adjusting something on oneself, like one's eyeglasses or clothing. Other-adapters entail rubbing an external object, i.e., something that is not on the person, such as the desk in the context of the present study.

4. Results

4.1. Quantitative results

4.1.1. Gestures with disfluencies

To answer the question of the role of gesture with disfluencies during SI, we first consider the relative amount of disfluencies that occurred with gesture phrases. Co-occurrence was assessed here with an ELAN search for temporal overlap, full or partial, between annotations of disfluency in the speech and annotations of gesture phrases. Considering the interpreting in both directions between Russian and English, 73% of the total amount of disfluencies in the data that were analyzed (950 of the total of 1300) occurred with gestures. In the Russian-German interpreting in both directions, 62% of the disfluencies in the data analyzed (579 instances out of 933) were produced with gestures. However, it is worth noting that the percent per individual interpreter varied greatly, namely from 10% to 85%.

In terms of the gesture functions with disfluencies, for the interpreting in both directions both between Russian and English and Russian and German, the gestures most commonly used were self-adapters or those serving pragmatic functions. This is indicated in Table 1.

	RUS-ENG & ENG-RUS		RUS-GER & GER-RUS	
	N	%	%	N
Self-adapter	415	**44%**	**54%**	313
Pragmatic	417	**44%**	**35%**	201
Representational	53	6%	5%	32
Deictic	23	2%	3%	17
Other-adapter	42	4%	3%	16
Totals:	950	100%	100%	579

Table 1. Gesture functions used with disfluencies

4.1.2. Gestures without disfluencies

The gestures that were produced in the 196 minutes analyzed that did not occur with disfluencies were also analyzed according to their functions (N=1250 in the Russian-English interpreting and 592 in the Russian-German interpreting, in both cases covering the interpreting in both directions). Once again, the predominant categories were self-adapters and pragmatic gestures, as shown in Table 2, but with pragmatic gestures predominating somewhat more as compared with the results of gestures with disfluencies.

	RUS-ENG & ENG-RUS		RUS-GER & GER-RUS	
	N	%	%	N
Self-adapter	477	**38%**	**43%**	255
Pragmatic	654	**52%**	**41%**	240
Representational	59	5%	8%	49
Deictic	37	3%	5%	30
Other-adapter	23	2%	3%	18
Totals:	1250	100%	100%	592

Table 2. Gesture functions used without disfluencies

4.1.3. Discussion of quantitative results

The similarly frequent use of self-adapters and pragmatic gestures, and infrequent use of representational and deictic gesture, during disfluency in rendering utterances and during non-disfluent interpreting suggests that manual gestures may not play a role that is unique to moments of disfluency. Rather, these categories of gesture may relate to functions of co-verbal behavior while interpreting in general, as discussed below. The stress that interpreters experience in their work is not something that is turned on like a switch during disfluencies and turned off during fluent interpreting. Instead, interpreters are constantly managing the cognitive load of the task, and while some moments involve a greater cognitive load than others (Chen, 2017), the stress is spread to varying degrees throughout the task (see, for example, Gile, 2008, on how a problem trigger in the source text might lead to a difficulty in rendering utterances not in the moment but further 'downstream' in the interpreter's flow of talk).

Considering the functions of adapters and pragmatic gestures, adapters are known to be related to self-regulation of stress (Ekman & Friesen, 1969; Freedman, 1972, 1977). In that regard, their use may help many simultaneous interpreters try to gain control over the interpreting process and decrease the level of cognitive load (Iriskhanova et al., 2019). While this may generally be the case, it is also worth bearing in mind that there was wide variation found across the individuals in this study in their degree of use of self-adapters, in line with the variation found in the use of gestures overall. This can be a factor of individual gesture styles—gestural idiolects—or as Lemmens (2015) calls them, idiogests.

The reason for such frequent use of pragmatic gestures may be less intuitively obvious. However, pragmatic gestures are known to help speakers structure and organize their discourse (Kendon, 2004, and many others), i.e., as a form of "speech-handling" (Streeck, 2009). Simultaneous interpreters are rendering not only referential content, but are also negotiating more abstract categories like information structure and stance—and gesture may participate in that process (Galhano-Rodrigues, 2007; Iriskhanova & Makoveyeva, 2020), on which see section 4.2.2 below.

The issue that remains is why the interpreters did not use many representational or deictic gestures. The low use of the former may relate to at least three factors: the cognitive processes behind the production of representational gestures, the time constraints of the process of SI, and the absence of any audience viewing the interpreters. The literature on representational gestures argues that their production may stem from mental simulation of the content that is being uttered verbally; this is Hostetter and Alibali's (2008, 2019) hypothesis of Gestures as Simulated Action. However, the cognitive process of SI is known to normally not involve any deeper semantic processing of the content of the speech being rendered than is needed to perform the interpreting (Alexieva, 1998; Riccardi, 1998). Thinking for SI is thus different from the kind of thinking for speaking involved in conversation, for example. Therefore, it does not involve the same kinds of growth points of ideas that McNeill (1992) considers as the sources of gestures with speech, particularly when it comes to representational gestures that relate to imagery associated with the content of the speech. This aligns with the findings in Leonteva et al. (2023) that abstract notions represented metaphorically in speakers' gestures were most often not carried over by interpreters viewing the speakers; when the interpreters did produce gestures in their renderings at similar points as the original speaker, they were most often pragmatic, presentation gestures, involving minimal metaphoricity (e.g., only schematically showing presentation of an idea with a relaxed hand, rather than tracing or molding more detailed imagery with a more tense hand). Furthermore, previous research (Alibali et al., 2001) found that speakers produced more representational gestures in a face-to-face condition than when listeners could not see their gestures, so this factor could also have come into play here.

The low use of deictic gestures could have to do with the lack of visual input that the interpreters had, only listening to the lectures rather than seeing the speakers, and the context of their working in an interpreting booth, looking into an empty classroom. The participants therefore had no inherent spatial grounding of the referents they were speaking about, nor any visually located deictic center of the speaker whom they were hearing. They also did not have any supporting visual aids to refer to that the original speaker might have used, such as slides being shown. In addition, as there was no one observing the interpreters, there was no interlocutor for whom the deixis would be needed.

4.2. Discussion of some qualitative findings

4.2.1. Self-adapters

We found great variation across individuals not only in how much they made use of self-adapters, but also in the manner in which they produced them. For example, a number of the interpreters had been trained in a tradition requiring them to keep their hands folded on the desk in front of them while interpreting. The logic behind that training is that if they should be visible to the listening audience, they should not be seen to be producing much visibly dynamic behavior, which could detract attention from the speaker of the source text. However, while a few of the participants did sit almost motionless at the desk during the task, others exploited the posture with hands folded to produce small self-adapters. Given that in this position the fingers of the interpreter's hands were often interlaced, the self-adapters sometimes involved micro-movements of tension and relaxation of the fingers, sometimes as one hand was gripping the other. In other cases, the movements were larger in amplitude, for example with the hands clasping and gripping each other in a more effortful fashion. Most of these self-adapters were sustained in nature over varying lengths of time.

4.2.2. Pragmatic functions of gestures

Gestures with pragmatic functions also ranged in terms of the specific functions they served and in the degrees of effort involved in their production, leading to lesser or greater salience. Many times, the pragmatic function involved was that of presenting an idea. While the palm-up open hand (as in Figure 1) is the gesture that has probably been researched the most as the gesture form serving this function (Bressem & Müller, 2014; Cooperrider et al., 2018; Müller, 2004), the position the interpreters often assumed with hands or arms folded on the desk in the booth afforded (and constrained) variations in how this was produced, as shown in Figures 2 and 3. Very often a simple turn-out (rotation outward) of the hand and upper arm was involved, as in Figure 2. Sometimes the mere raising of a finger served the same function in a very small fashion, as an outward beat emphasizing a point being made in the speech. With the hands folded, this sometimes just took the form of one or both thumbs being extended upward and then lowered, as in Figure 3. Cienki (2021) discusses these as a continuum of pragmatic gestures, ranging from a finger-lift, to a rotation outward of the hand and upper arm, to a full extension outward of a palm-up open hand.

Figure 1. A (double) palm-up open hand pragmatic gesture when presenting an idea

Figure 2. A turn-out of the hand when presenting an idea

Figure 3. A thumb extension when presenting an idea

In other cases, the pragmatic function was one of more distinct stance-taking. For example, the lifting of one or both shoulders and/or a head tilt or head shake, sometimes accompanied by the opening and turning out of one or both hands, comprise elements of a shrug (Debras & Cienki, 2012). This can reflect a range of stances from indicating uncertainty, to incredulity, to distancing oneself from another's views on a topic (Debras, 2017). In one instance, the interpreter uttered the words in Russian, "*Èto poterya antropotsena. Poteri kolossal'ny.*" ('This is a loss from the Anthropocene. The losses have been colossal.'), and when saying 'colossal' he lifted his right shoulder and quickly shook his head. The co-verbal behavior suggests that the amount of the losses is unbelievably large. Again, it is important to remember that the interpreter was not viewing any video of the original speaker; the gesture was of his own creation. In other instances, some interpreters gestured with the tips of the thumb and index finger pressed together as they mentioned a specific number, highlighting the exactness of the amount with what is known as a precision grip (Kendon, 2004, ch. 12). Such instances raise interesting questions about whose stance the interpreter is expressing (their own, or the imagined stance of the original speaker), and whether this can even be ascertained—something considered further in Cienki (2024).

4.2.3. Other functions of gestures

As described above, the interpreters made little use of representational or deictic gestures. Many of the representational gestures that did occur involved the holding mode of representation when mentioning a quantity (e.g., "two point five species") or a fact (e.g., "what happens to these species"); see Figure 4.

Figure 4. Holding gesture when mentioning a quantity
(here namely: "two point five species")

The kind of representation involved in Figure 4 is quite schematic, whereby the interpreter is as if holding the amount being mentioned, with fingers spread and slightly curved, the palm of each hand turned toward the space in front of her. In this regard, even in many of the gestures with a primarily representational function, one could see a secondary pragmatic function similar to that of presenting an idea with a hand turn-out or palm-up open hand.

Deictic pointing to spaces was rarely used, but the few instances in which it did occur present interesting phenomena. In one case, the speaker of the source text said in German, "*Hier aufgetragen die Summe der Brutreviere in den erfassten Quadranten rund um ähm den Bodensee. Sie sehen, dass es Verluste – rot – und natürlich auch Arten gibt, die...*" ('Here the sum of the breeding territories is plotted in the quadrants recorded around, um, Lake Constance. You can see that there are losses – red – and of course there are species which...'). The interpreter rendered this in Russian as "*eh vy vidite to, chto dannye poteri, oni oboznacheny krasnym na skheme*" ('uh, you can see that the data on the losses, they are indicated in red on the chart') and on the words 'data' and 'indicated' he pointed to the upper right and he looked up to the right during that entire stretch of speech. In this instance, we see the interpreter presenting an imagined deictic viewpoint of the speaker of the source text. Interestingly, the original lecturer had actually shown the chart on his left side and did not point to it when he made reference to it. The deictic reference in gesture and eye gaze was completely the interpreter's invention.

5. Closing points

The findings from the present study show quantitatively similar patterns of use of gesture functions both during moments of disfluency in interpreted speech and during fluent interpreting. Rather than highlighting a special function for gesture during disfluency in SI,

the results suggest more general, overarching roles that gesture plays in this context. The results, with pragmatic gestures and self-adapters having been by far the most frequently used functions, stand in contrast to findings by, for example, McNeill (1992), who found representational gestures (his categories of iconic and metaphoric gestures) to be used even more than pragmatic gestures (his category of beat gestures) in narratives[6]. Overall, this points to the potentially different kind of thinking that is involved in speaking for SI than is normally involved in thinking for speaking (à la Slobin, 1987) in self-initiated talk, as in conversation or unrehearsed narratives. The fact that representational gestures played such a small role in the interpreters' gestural repertoire might be a reflection of not engaging in the unpacking of idea units ("growth points") in the way that McNeill described, but of converting ideas received via one language into another language, and mostly not engaging in deep semantic processing, as Alexieva (1998) and Riccardi (1998) argue. SI is known to entail specialized forms of cognitive processing (García, 2019) and so it makes sense that gesture during SI of a lecture would differ from gesture use during another form of monologic speech, namely self-initiated narration, given the relation of gesture to conceptualization (Kita et al., 2017). In this regard, it is perhaps ironic that it was McNeill's observation of a simultaneous interpreter's gestures that sparked his interest in the relations between thought, spontaneous speech, and gesture.

The interpreters' extensive use of pragmatic gestures and sustained self-adapters highlights two aspects of their role in performing this work. On the one hand, they are presenters of another's ideas to an audience in a different language. In this respect, the use of pragmatic gestures is logical, given the role they are known to play in interaction. Such gestures are outwardly oriented, not only in their form, moving out from the speaker's body, but also in their functions, such as presenting ideas to others for their consideration, or showing one's stance towards the ideas being presented. The interpreters engaged with this role as part of their practice, even when sitting alone in an interpreting booth with no other speaker or listener in view[7]. On the other hand, simultaneous interpreters are dealing with a heavy cognitive load as part of their work. Self-adapter movements may help them handle this through the combined effects that body-focused movements can have of assisting in maintaining one's mental focus while also soothing oneself during stress (Ekman & Friesen, 1969; Freedman, 1972). In this way, self-adapters can be seen as a reflection of the inwardly oriented cognitive and affective aspects that are part of SI.

As argued in Cienki and Iriskhanova (2020), simultaneous interpreters blend the viewpoint of themselves as speakers with the imagined or perceived viewpoint of the speaker of the source text. The fact that interpreters' co-verbal behaviors were found to be generally similar during moments of disfluency and during fluent interpreting suggests that the combining of inward- and outward-oriented perspectives is a process being negotiated throughout the process of interpreting.

6. Acknowledgments

The research was supported by Russian Science Foundation grant number 19-18-00357 for the project "Verbal and co-verbal behavior under cognitive pressure: Analyses of speech, gesture, and eye gaze" carried out at Moscow State Linguistic University. Credit for the data collection, annotation, and analysis is due to the following members of the PoliMod Lab at the Centre for Socio-Cognitive Discourse Studies (SCoDis); in alphabetical order: Olga Agafonova, Olga Iriskhanova, Varvara Kharitonova, Anna Leonteva, Alina Makoveyeva, Andrey Petrov, Olga

[6] Iconic = 261 instances, metaphoric = 43, beat = 268 (McNeill, 1992, p. 93).

[7] Though they knew they were being videorecorded, the participants rarely looked directly at the small GoPro camera on the desk in front of them, but rather looked into various spaces in the booth and in the classroom in front of them, and sometimes closed their eyes.

Prokofeva, Evgeniya Smirnova, and Maria Tomskaya, as well as Geert Brône.

7. References

Ahrens, B. (2007). Pauses (and other prosodic features) in simultaneous interpreting. *FORUM. Revue internationale d'interprétation et de traduction / International Journal of Interpretation and Translation*, *5*(1), 1–18. https://doi.org/10.1075/forum.5.1.01ahr

Alexieva, B. (1998). Consecutive interpreting as a decision process. In A. Beylard-Ozeroff, J. Kralova, & B. Moser-Mercer (Eds.), *Translators' strategies and creativity* (pp. 181–188). John Benjamins.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production. *Journal of Memory and Language, 44*(2), 169–188. https://doi.org/10.1006/jmla.2000.2752

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*, 469–489. https://doi.org/10.1080/01638539209544823

Bressem, J., & Müller, C. (2014). A repertoire of recurrent gestures of German with pragmatic functions. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & J. Bressem (Eds.), *Body – language – communication*, Volume 2 (pp. 1575–1591). De Gruyter Mouton.

Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives*, *25*(4), 640–657. https://doi.org/10.1080/0907676X.2016.1278026

Cienki, A. (2021). From the finger lift to the palm-up open hand when presenting a point. *Languages and Modalities*, *1*, 17–30. https://doi.org/10.3897/lamo.1.68914

Cienki, A. (2024). Variable embodiment of stance-taking and footing in simultaneous interpreting. *Frontiers in Psychology, 15*, Article 1429232. https://doi.org/10.3389/fpsyg.2024.1429232

Cienki, A., & Iriskhanova, O. K. (2020). Patterns of multimodal behavior under cognitive load: An analysis of simultaneous interpretation from L2 to L1. *Voprosy Kognitivnoy Lingvistiki*, *1*, 5–11. https://doi.org/10.20916/1812-3228-2020-1-5-11

Clift, R. (2020). Stability and visibility in embodiment: The 'Palm Up' in interaction. *Journal of Pragmatics*, *169*, 190–205. https://doi.org/10.1016/j.pragma.2020.09.005

Cooperrider, K., Abner, N., & Goldin-Meadow, S. (2018). The palm-up puzzle. *Frontiers in Communication, 3*, Article 23. https://doi.org/10.3389/fcomm.2018.00023

Dayter, D. (2020). Variation in non-fluencies in a corpus of simultaneous interpreting vs. non- interpreted English. *Perspectives*, *29*(4), 489–506. https://doi.org/10.1080/0907676X.2020.1718170

Debras, C. (2017). The shrug: Forms and meanings of a compound enactment. *Gesture*, *16*(1), 1–34. https://doi.org/10.1075/gest.16.1.01deb

Debras, C., & Cienki, A. (2012). Some uses of head tilts and shoulder shrugs during human interaction, and their relation to stancetaking. *International Proceedings of the ASE 2012 International Conference of Social Computing*. Amsterdam, Netherlands, 3–5 September 2012. https://doi.org/10.1109/SocialCom-PASSAT.2012.136

Dressel, D. (2020). Multimodal word searches in collaborative storytelling. *Journal of Pragmatics*, *170*, 37–54. https://doi.org/10.1016/j.pragma.2020.08.010

Du Bois, J. D., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data* (pp. 45–89). Lawrence Erlbaum Associates.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior. *Semiotica*, *1*(1), 49–98. https://doi.org/10.1515/semi.1969.1.1.49

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. Wolfe Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 153–175). Pergamon Press.

Freedman, N. (1977). Hands, words, and mind. In N. Freedman & S. Grand (Eds.), *Communicative structures and psychic structures* (pp. 109–132). Springer.

Galhano-Rodrigues, I. (2007). Body in interpretation. In P. Schmitt & H. Jüngst (Eds.), *Translationsqualität. Leipziger Studien zur angewandten Linguistik und Translatologie* (pp. 739–753). Peter Lang Verlag.

García, A. M. (2019). *The neurocognition of translation and interpreting.* John Benjamins.

Gibbon, D. (2005). Prerequisites for a multimodal semantics of gesture and prosody. *Proceedings of the International Workshop on Computational Semantics*, *6*, 2–15.

Gile D. (1997), Conference interpreting as a cognitive management problem. In J. H. Danks, S. B. Fountain, M. K. McBeath, & G. M. Shreve (Eds.), *Cognitive processes in translation and interpreting* (pp. 196–214). Sage Publications.

Gile, D. (2008). Local cognitive load in simultaneous interpreting and its implications for empirical research. *Forum*, *6*(2), 59–77. https://doi.org/10.1075/forum.6.2.04gil

Gósy, M. (2007). Disfluencies and self-monitoring. *Govor = Speech*, *24*(2), 91–110. https://hrcak.srce.hr/clanak/256310

Grishina, E. A. (2017). *Russkaya zhestikulyaciya s lingvisticheskoy tochki zreniya* [Russian gesticulation from a linguistic point of view]. Yazyki slavyanskoy kul'tury.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514. https://doi.org/10.3758/PBR.15.3.495

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action? Revisiting the framework. *Psychonomic Bulletin & Review*, *26*, 721–752. https://doi.org/10.3758/s13423-018-1548-0

Iriskhanova, O. K., & Makoveyeva, A. I. (2020). O pragmaticheskoy funkcii zhestov v sinkhronnom perevode [On the pragmatic function of gestures in simultaneous interpreting]. *Kognitivnye issledovaniya yazyka [Cognitive Studies on Language]*, *3*(42), 54–60. https://www.elibrary.ru/item.asp?id=44231536

Iriskhanova, O. K., Petrov, A. A., Makoveyeva, A. I., & Leonteva, A. V. (2019). Kognitivnaya nagruzka v usloviyakh sinkhronnogo perevoda [Cognitive load in the context of simultaneous interpreting]. *Kognitivnye issledovaniya yazyka [Cognitive Studies on Language]*, *38*, 100–115. https://www.elibrary.ru/item.asp?id=41350283

Kajzer-Wietrzny, M., Ivaska, I., & Ferraresi, A. (2024). Fluency in rendering numbers in simultaneous interpreting. *Interpreting*, *26*(1), 1–23. https://doi.org/10.1075/intp.00092.kaj

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relation of verbal and nonverbal communication* (pp. 207–227). Mouton.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kendon, A. (2016). Gesture and sign: Utterance uses of visible bodily action. In K. Allan (Ed.), *The Routledge handbook of linguistics* (pp. 33–46). Routledge.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and* gesture (pp. 162–185). Cambridge University Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245–266. https://doi.org/10.1037/rev0000059

Kok, K., Bergmann, K., Cienki, A., & Kopp, S. (2015). Mapping out the multifunctionality of speakers' gestures. *Gesture*, *15*, 37–59. https://doi.org/10.1075/gest.15.1.02kok

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge University Press.

Ladewig, S. H. (2011). Putting the cyclic gesture on a cognitive basis. *CogniTextes*, 6, article 406. https://doi.org/10.4000/cognitextes.406

Lemmens, M. (2015). Idiogests: Gestural interspeaker variation reveals semantic variation. Presentation at AFLiCo 6, Grenoble, France, 26-28 May, 2015.

Leonteva, A. V., Cienki, A., & Agafonova, O. V. (2023). Metaphoric gestures in simultaneous interpreting. *Russian Journal of Linguistics*, *27*(4), 820–842. https://doi.org/10.22363/2687-0088-36189

Martín de León, C., & Fernández Santana, A. (2021). Embodied cognition in the booth: Referential and pragmatic gestures in simultaneous interpreting. *Cognitive Linguistic Studies*, *8*(2), 277–306. https://doi.org/10.1075/cogls.00079.mar

Mazza, C. (2001). Numbers in simultaneous interpreting. *The Interpreters' Newsletter*, *11*, 87–104.

McNeill, D. (1992). *Hand and mind*. University of Chicago Press.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.

McNeill, D., Cassell, J., & Levy, E. T. (1993). Abstract deixis. *Semiotica, 95*(1/2), 5–19. https://doi.org/10.1515/semi.1993.95.1-2.5

McNeill, D., & Duncan, S. D. (2000). Growth points in thinking for speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge University Press.

Müller, C. (1998a). Iconicity and gesture. In S. Santi, I. Guaïtella, C. Cave, & G. Konopczynski, (Eds.), *Oralité et gestualité: Communication multimodale et interaction* (pp. 321–328). L'Harmattan.

Müller, C. (1998b). *Redebegleitende Gesten*. Berlin Verlag A. Spitz.

Müller, C. (2004). Forms and uses of the Palm Up Open Hand. In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gestures* (pp. 233–256). Weidler.

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & J. Bressem (eds.), *Body – language – communication* (pp. 1687–1702). De Gruyter Mouton.

Pellatt, V. (2006). The trouble with numbers. In M. Chai & J. Zhang (Eds.), *Professionalization in interpreting* (pp. 350–365). Foreign Language Education Press.

Pinochi, D. (2010). Simultaneous interpretation of numbers: Comparing German and English to Italian. *The Interpreters' Newsletter*, *14*, 33–57.

Plevoets, K., & Defrancq, B. (2016). The effect of informational load on disfluencies in interpreting. *Translation and Interpreting Studies*, *11*(2), 202–224. https://doi.org/10.1075/tis.11.2.04ple

Plevoets, K., & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament. *Interpreting*, *2*(1), 1–28. https://doi.org/10.1075/intp.00001.ple

Ren, Y., & Wang, J. Gesture and cognitive load in simultaneous interpreting: A pilot study. *Parallèles, 37*(1), 47-65. https://doi.org/10.17462/PARA.2025.01.04

Riccardi, A. (1998). Interpreting strategies and creativity. In A. Beylard-Ozeroff, J. Kralova, & B. Moser-Mercer (Eds.), *Translators' strategies and creativity* (pp. 171–179). John Benjamins.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Seeber K. G. (2013). Cognitive load in simultaneous interpreting. *Target*, *25*(1), 18–32. https://doi.org/10.1075/target.25.1.03see

Slobin, D. I. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society* (pp. 435-445). Berkeley Linguistics Society. https://doi.org/10.3765/bls.v13i0.1826

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)* (pp. 816–820). https://hdl.handle.net/11858/00-001M-0000-0013-1F92-C

Streeck, J. 2009. *Gesturecraft: The manu-facture of meaning.* John Benjamins.

Vygotsky, L. (1962). *Thought and language.* MIT Press. (Original work published 1934 as *Myshlenie i rech'* [Thinking and speech]. Labirint)

Zagar Galvão, E. (2015). *Gesture in simultaneous interpreting from English into European Portuguese*. PhD dissertation, University of Porto, Portugal.

Zagar Galvão, E. (2020). Gesture functions and gestural style in simultaneous interpreting. In H. Salaets & G. Brône (Eds.). *Linking up with video* (pp. 151–179). John Benjamins.

Zagar Galvão, E., & Galhano-Rodrigues, I. (2010). The importance of listening with one's eyes. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.), *New insights into audiovisual translation and media accessibility* (pp. 241–253). Rodopi.

Alan Cienki
Vrije Universiteit Amsterdam
FGW, postvak 4.14, De Boelelaan 1105
1081 HV Amsterdam
Netherlands
a.cienki@vu.nl

Biography: Alan Cienki is Professor of Language Use & Cognition and English Linguistics at VU Amsterdam where he directs the MA program in Multimodal Communication. His research in cognitive linguistics focuses on the interplay between spoken language and gesture in expressing meaning. He edited *The Cambridge handbook of gesture studies* (2024) and co-edited the two-volume handbook *Body–language–communication* (2013, 2014). He has served as Chair of the association for Researching and Applying Metaphor and Vice President of the International Society for Gesture Studies. In 2013 he founded the Multimodal Communication and Cognition Lab "PoliMod" at Moscow State Linguistic University.

Parallèles 37(1) April 2025

Gesture and cognitive load in simultaneous interpreting: A pilot study

Yuetao Ren

University of International Business and Economics (China)

Jianhua Wang

Renmin University of China

Abstract

This paper explores the relationship between gesture and cognitive load in simultaneous interpreting (SI). To this end, we set up a remote interpreting setting for data collection. Thirteen master's student interpreters participated in two SI tasks, one in a video condition and the other in an audio condition. We analyzed their gestural behaviors and disfluency patterns, as well as the correlation and temporal relation between gestures and disfluencies. We found that interpreters gestured more in tasks with a higher cognitive load (audio interpreting), although the differences in disfluency rate and gesture rate between the two conditions were not significant. Even though the correlation between gesture and cognitive load was not significant, all the gestures in the study were produced parallel with or adjacent to processing difficulties. We conclude that gestures could be the embodied manifestation of the cognitive processes of SI and of the 'exported load'. Furthermore, the function of each gesture type varies under cognitive load. Silent gestures (beats and metaphorics) may reflect the interpreter's use of strategies, while the production of semantically related gestures (deictics and iconics) may be influenced by cognitive load. The results contribute to the understanding of SI as an embodied, multimodal cognitive activity.

Keywords

Simultaneous interpreting, gesture, cognitive load, disfluency, multimodality

 DOI:10.17462/para.2025.01.04

1. Introduction

Interpreting is generally acknowledged as a cognitively demanding activity (Chen, 2017). Simultaneous interpreting (SI), one of the major working modes of interpreting in conference settings, is considered the most complex language task (Christoffels & De Groot, 2005). SI is often studied from a cognitive perspective, focusing on its multitasking nature and high cognitive demand.

In recent years, the idea that cognition is embodied, embedded, extended, enactive, and affective has been adopted in the cognitive study of translation and interpreting (Muñoz Martín, 2016). In this paper, we adopt the embodied approach to cognition, which holds that cognitive processes are deeply rooted in the body's interactions with the world (Wilson, 2002). Embodiment explains thinking and speaking as perceptually and motorically based (Barsalou, 2008). The representation and manipulation of information is accomplished via the simulation of sensorimotor processes, and these simulation mechanisms may give rise to both speech and gesture (Hostetter & Alibali, 2008). When performing a complex language task like SI, the interpreter employs a wide range of embodied modalities like gesture, gaze, and posture. Among these modalities, gestures are intimately connected with language and have been shown to perform self-oriented functions, facilitating the cognitive processes of thinking and speaking (Kita *et al*., 2017).

Studies on the multimodal behavior of interpreters have revealed that interpreters do gesture during SI, even when they are 'invisible' in the booth (Adam & Castro, 2013; Cienki, this volume; Cienki & Iriskhanova, 2020; Martín de León & Fernández Santana, 2021; Martín de León & Zagar Galvão, this volume; Zagar Galvão, 2020). However, a systematic study of the embodied, multimodal cognitive processes of SI is still lacking. Gestures are outward manifestations of the cognitive processes that govern thinking and speaking. They open a "window onto the mind" (McNeill, 1992, p. 268). As such, gestures may provide visual clues for the study of the cognitive processes of SI. This study is a pilot for such research. We include the gestural behaviors of simultaneous interpreters in our analysis while focusing on one of the most distinctive aspects of SI, i.e. cognitive load. Before introducing the methodology, we will review the theoretical aspects of cognitive load in SI, discuss the role of gestures in thinking and speaking, and comment on the existing studies of gesture and cognitive processes of SI.

1.1. Cognitive load in simultaneous interpreting

Following Seeber (2013, p. 19), we define cognitive load as "the amount of capacity the performance of a cognitive task occupies in an inherently capacity-limited system." This definition is based on the assumption that working memory is a capacity-limited system (Seeber, 2011). In other words, the number of operations the human brain can carry out and the amount of information it can maintain for processing at a given time is limited. There are two models that offer explicit accounts for cognitive load in interpreting: Effort Models (Gile, 1999, 2008) and the Cognitive Load Model (Seeber, 2011; Seeber & Kerzel, 2012).

Effort Models are a set of models that account for operational constraints in interpreting (Gile, 1999). Interpreting is conceptualized as a set of multiple cognitive operations which can be grouped into certain 'Efforts', such as listening and analysis (L), speech production (P), and short-term memory (M) (Gile, 2008). Effort Models assume a single pool of resources for all efforts. It is further assumed that different efforts may compete for the total available processing capacity ('Competition Hypothesis') and interpreters tend to work close to saturation ('Tightrope Hypothesis') (Gile, 1999). Problem triggers, such as numbers, lists, and proper names, are associated with the increase of processing capacity requirements, which may exceed the total amount of capacity available or cause energy management problems,

resulting in errors, omissions, and/or reduced quality in the performance. Furthermore, the load might be carried over into downstream segments ('exported load') (Gile, 2008). Problem triggers do not necessarily lead to actual problems in the corresponding segments but may affect other segments, which are at a distance and are not difficult to render in and of themselves.

The Cognitive Load Model accounts for the effect of different combinations of sub-tasks on overall cognitive demands (Seeber, 2011). The load is not generated from the competition of resources, but rather from the interference between sub-tasks. The Cognitive Load Model distinguishes different processing codes (verbal-spatial) and modalities (auditory-visual). The former refers to the different systems of working memory, and the latter refers to the sensory modalities of input and response. According to this model, tasks of the common structures interfere with each other more strongly than those of different structures. In other words, processing codes and modalities have an effect on task interference. Discrete spatial and verbal tasks are time-shared more efficiently than two spatial or two verbal tasks, and intra-modal processes interfere with each other more than inter-modal processes (Seeber, 2007).

Effort Models and the Cognitive Load Model share some common ground. As Seeber (2011) argues, the two models aim to account for the cognitive demands inherent to interpreting. They both see interpreting as multitasking, which is comprised of a set of sub-tasks. Furthermore, Seeber and Kerzel (2012) present empirical evidence that corroborates Gile's (2008) idea of exported load, as the relative maximum local cognitive load occurs at the end of the sentence.

1.2. The role of gesture in thinking and speaking

1.2.1 Classifications of gesture

Co-speech gestures are taken here to be spontaneous speech accompaniments that are made with fingers, hands, and arms (McNeill, 2005). Gestures can be classified into several subtypes. McNeill (1992) distinguished four types in terms of the forms and functions of gestures: iconics, metaphorics, beats, and deictics. Iconic gestures represent concrete concepts by depicting the shape, size, or contour of the referent. Metaphoric gestures are similar to iconics in form but are associated with abstract concepts. Beats are biphasic movements of the finger or hand, which can serve an emphatic function. Deictic gestures are pointing movements that refer to an entity or a space by extending the finger, hand, or arm.

Kendon (2004) outlined three main ways in which gestures contribute to the meanings of utterances, including via referential, pragmatic, and interactional functions. Referential gestures contribute to the propositional meanings of the utterances. Pragmatic gestures contribute to the acts accomplished by utterances, such as indicating the speaker's attitude, providing an interpretative framework, or making manifest the speech act. Interactional gestures are used to regulate interactions. The referential and pragmatic functions of gestures are not mutually exclusive. For example, pointing gestures mainly serve referential functions, but the different hand shapes used in pointing may have pragmatic functions.

In this paper, we adopt McNeill's (1992) categorization of gestures, which is based on a psycholinguistic perspective, and gestures were perceived as a "window onto the mind" (p. 268).

1.2.2 Gesture-speech integration

Gesture and speech form an integrated system during language production (McNeill, 1992, 2005). The minimal idea unit is a 'growth point' (GP), which consists of both imagery and linguistic content and can develop into a full utterance with a gesture (McNeill, 1992, 2005). In other words, gesture and language share the same computational stage (McNeill, 1985), and

the product of this stage is a concept that can be packaged and expressed both in speech and in gesture.

Gesture either synchronizes with a parallel linguistic unit or comes *before* the linguistic expression, suggesting that gesture can reveal the moment at which the speaker formulates a concept (McNeill, 1985). McNeill distinguished two kinds of 'gestural anticipation' (p. 361). First, during uninterrupted speech, semantic computation takes place and is expressed in gesture, while the corresponding linguistic expression for the same concept may be delayed; it comes *after* the linguistic segment that goes with the gesture. Such gestures include iconic and deictic gestures; they refer to the content of the utterance and perform a referential function (Kendon, 2004). Second, during silence, where speech comes to a halt, there is no semantic computation taking place. Beat or metaphoric gestures may be produced during silence. Such 'silent gestures' are part of speaking; they provide a metalinguistic commentary on the process of speaking (McNeill, 1985, p. 354), fulfilling a pragmatic function (Kendon, 2004).

Gesture is involved in cognitive processes by activating, manipulating, packaging, and exploring information for thinking and speaking (Kita *et al*., 2017). Gestures have been argued to be generated from the same process that generates practical actions (Hostetter & Alibali, 2008). Thus, gestures can influence thoughts about both spatio-motoric information based on bodily experiences, and about abstract information, via the metaphorical use of spatio-motoric information. However, gestures are different from practical actions in that they are schematic representations. Focusing on essentials rather than details, such representations can be processed efficiently and are flexible and modifiable (Kita *et al*., 2017). As such, gestures play a facilitative role in the cognitive processes of thinking and speaking.

1.3. Gestures and the cognitive processes of interpreting

Recently, interpreting has been conceived as a multilingual and multimodal embodied cognitive activity (Cienki & Iriskhanova, 2020; Martín de León & Fernández Santana, 2021; Stachowiak-Szymczak, 2019). The input interpreters receive is inherently multimodal, including both auditory and visual information. When producing the output, interpreters employ a range of embodied modalities and resources. One of the most essential and inherent resources for interpreters is gesture. In conference interpreting, researchers have focused on the cognitive aspects of the interpreter's gestures.

Adam and Castro (2013) investigated the form and function of beat gestures produced by student interpreters during SI. Results showed that beats were the most prevalent (84.7%) gestures produced by interpreters, while 18.41% of all the gestures appeared at moments of hesitation. When gestures appeared during pauses, interpreters were having comprehension problems or engaging in word searches. These gestures could have been used as an unconscious strategy by participants. When gestures accompanied self-corrections, they usually went with the corrected version of the utterance for emphasis.

Stachowiak-Szymczak (2019) also focused on interpreters' beat gestures in both simultaneous and consecutive interpreting (CI). Using numbers and lists as problem triggers, the study tested different levels of cognitive effort and correlated them with the gestural behaviors of the interpreter. Results showed that cognitive effort was reflected in gesture numbers. Both professional and student interpreters produced more gestures when interpreting lists and numbers compared to interpreting narratives. Beat gestures could have been produced by participants to deal with local cognitive effort in interpreting, especially for reducing the load related to processing lists.

Martín de León and Fernández Santana (2021) analyzed the interpreting process of one professional simultaneous interpreter. The study revealed the different roles of the interpreter's

gestures in the interpreting process. Using referential gestures may help the interpreter with source language (SL) comprehension, while producing pragmatic gestures could lend support to her target language (TL) production.

Taken together, results of the above studies show that gestures employed by interpreters are related to the cognitive processes of interpreting, especially at moments when cognitive demands are high. However, systematic analysis of gestures and cognitive processes is lacking. Stachowiak-Szymczak (2019) focused on only one type of gesture, while Martín de León and Fernández Santana (2021) studied only one interpreter. In this study, we try to expand the scope of previous studies by including all the co-speech gestures produced by eleven interpreters.

This study aims to explore the relationship between the interpreter's gestural behavior and cognitive load in simultaneous interpreting. It was guided by the following research questions:

1) Is there a correlation between the interpreter's gesture and cognitive load in simultaneous interpreting?

2) If so, what functions do gestures of different types play under cognitive load?

For the first question, we expect that more gestures are produced under high cognitive load. Following Gumul (2021), we used disfluencies as indicators of cognitive load. The assumption is that disfluencies are evidence of a decrease in interpreting quality, which is likely to be associated with an increase in cognitive load (Chen, 2017, p. 647). In other words, we expect that gestures are produced more with disfluent speech than with fluent speech (Cienki, this volume).

For the second question, we formulated two hypotheses about the functions of gestures of different types under cognitive load. Based on McNeill's (1985) notion of gestural anticipation, we categorized gestural behaviors under cognitive load into two kinds. Silent gestures are associated with processing difficulty, as they often arise during speech breakdown. In contrast, gestures produced alongside uninterrupted speech might not be accompanied by processing difficulties. Following McNeill's (1992) categorization of gesture types, we expect that all gesture types, including iconics, metaphorics, beats, and deictics, are produced in the interpreting process; but only beat and metaphoric gestures are produced with processing difficulty (silent gesture). Moreover, we expect that deictic and iconic gestures are produced before the production of their linguistic counterparts (gestural anticipation), without processing difficulty.

2. Methodology

2.1. Participants

Thirteen Chinese students in a Master of Translation and Interpreting program (MTI) (12 females and 1 male) participated in the study on a voluntary basis. The average age was 23.9 years (*SD* = 1.18 years, range 22 – 26 years). They had the same language combination, with Mandarin Chinese as L1 and English as L2. They all have passed the English language test TEM-8[1], with an average score of 75 (*SD* = 3.25, range 71 – 81). None of them had worked as a professional interpreter before. By the time of the experiment, they had received CI training for two semesters and SI training for one semester. Two participants (both females) did not perform any co-speech gestures in any of the interpreting tasks, but they did use such gestures in other speech production tasks in the experiment. Since our focus was co-speech gestures in SI, we excluded them from the data analysis for this study. The final *N* here was 11 (10 females and 1 male).

[1] TEM-8, which stands for Test for English Majors Band 8, is a Chinese equivalent of IELTS. Its full mark is 100.

All participants were tested to be right-handed using the Edinburgh Inventory Handedness Questionnaire[2]. They were told that the whole experiment would be videotaped and they signed a written consent before the experiment. After the experiment, they were given another consent form which concerned their willingness to have their video images used in academic reports. To protect the anonymity and confidentiality of participants, they were informed that their faces would be blocked when using their images. All participants agreed to participate in the experiment and to have their images used anonymously. They were given 100 RMB compensation for their time and efforts in participating in the experiment.[3]

2.2. Materials

We selected two speech videos of different topics from the same speaker. Speech A was adapted from an extended talk[4]. We selected the first five minutes and a half starting from the beginning, in which the speaker illustrated her first point of the whole talk. Speech B is a complete TED talk[5].

We calculated the text complexity of the spoken texts using the Flesch Kincaid Readability Index[6], considering the reading ease (A: 72.7, B: 70.2) and the percentage of complex words (A: 10.67%, B: 10.20%). The Flesch Reading Ease score ranges from 1 to 100. The two texts correspond to a school grade level 8 (ages 12 to 14) and should be fairly easy for the average adult to read. Besides, the percentage of complex words in the two texts is close to one another, reflecting a similar complexity of the two texts.

In terms of speed, the original rate of speaking in speech A and B were similar (145.41 wpm vs. 145.19 wpm respectively). For SI, the speed of the source speech is a potential 'problem trigger' (Gile, 2008). Since participants were not professional interpreters and had only mastered basic skills for SI, fast speed would pose a challenge for them. We made some minor adjustments to slow down the original speed.

We randomly selected speech A and converted it into an audio file, while the manipulated speed remained unchanged. Thus, we produced two source speech conditions, namely one as a video and one as an audio-only recording. For the video speech (speech B), participants would hear and only see the speaker's image in the video, without captions, PPT slides, and other images. The rationale for having two interpreting conditions is the potential relevance of codes and modalities in cognitive load (Seeber, 2007, 2011). The comparison between the two conditions could unveil such potential effects.

The word count, final length, and speed of the materials, as well as corresponding interpreting conditions, are shown in Table 1.

	Source speech	
	Speech A	**Speech B**
Word count	807	861
Final length	6'10''	6'35''
Final speed	130.79 wpm	130.85 wpm
Interpreting condition	Audio	Video

Table 1. Summary of information about the materials for the experiment

[2] Available at http://www.brainmapping.org/shared/Edinburgh.php#

[3] The experiment was approved by the research ethics committee of Renmin University.

[4] https://www.youtube.com/watch?v=tBRjmjsbrcE

[5] https://www.ted.com/talks/angela_lee_duckworth_grit_the_power_of_passion_and_perseverance

[6] https://www.webfx.com/tools/read-able/

2.3. Procedure

We set up a simulated remote interpreting setting via Tencent Meeting[7], a Chinese equivalent of ZOOM, for data collection. Participants and the experimenter were in two separate rooms (partly as a result of Covid restrictions). They were connected via Tencent Meeting on two laptops. The experimenter played the speech files on his laptop, while the screen and sound were shared via Tencent Meeting and simultaneously displayed on the participant's laptop, which was placed on the right front of them. For the audio speech (speech A), participants would see a blank screen. They could not see the experimenter or other audience in either of the interpreting tasks. The whole process of the experiment was filmed by two cameras simultaneously from different perspectives. This design was intended to ensure that all hand movements, including finger movements, could be captured without obstruction.

In the warm-up part of the experiment, participants first talked freely in Chinese about their English learning experience. In this part, participants could get familiar with the remote interpreting condition, and the experimenter could check the state of the apparatus and internet connection. Then, participants did simultaneous interpreting in two different conditions. The order of the two sessions was randomized per participant. After each interpreting task, participants conducted a cued retrospection using their own videos as stimuli, in which they made spoken commentaries on their interpreting process. Participants could take a break of two to three minutes after the retrospection. Data were originally collected for the first author's PhD thesis from May to October 2022. Aside from the two SI tasks, participants also performed two CI tasks using different materials, each followed by a retrospection. The order of the four interpreting tasks was randomized for each participant. After interpreting, we conducted a semi-structured interview on the use of gestures in communication. For this study, we only focused on the two SI tasks. The CI data will be used for another study and are not presented here.

Twenty-four hours before the experiment, we gave participants some term lists, including proper names and unfamiliar jargon, with corresponding Chinese equivalents. Participants could preview the terms as preparation for the interpreting tasks. Before each interpreting task, they could review the terms for that specific task, but they could not refer to them during the process of interpreting.

We did not inform participants of the actual purpose before the experiment, because knowing this might potentially affect the spontaneity of their gestural behaviors. All they knew was that it would involve a remote interpreting experiment. It was only at the end of the experiment, during the debriefing, that participants got to know that their gestures were being studied. The whole session took about 100 minutes for each participant.

2.4. Data analysis

We focused on two types of product data for the present study, including gestures and disfluencies. We used the ELAN software (version 6.2) (Sloetjes & Wittenburg, 2008) for the annotation and analysis of data.

For disfluencies, we included the following sub-types: silent pause, filled pause, and false start. According to Han and An (2020), we chose 0.5s as the threshold for identifying silent pauses. Following Bóna and Bakti (2020), we defined a filled pause as a sound or syllable that does not contribute to the meaning of the sentence, like *uh* or *um* in English and '嗯' (*enn*) or '呃' (*err*) in Chinese. In this study, we used 'false start' as an umbrella term, which included restarts and truncation (Cienki & Iriskhanova, 2020), partial or whole word repetition, revision, broken words, and prolonged sounds (Bóna & Bakti, 2020).

[7] https://meeting.tencent.com/

For gestures, we focused on co-speech gestures, i.e. hand movements that had an intimate relationship with the co-occurring speech (McNeill, 2005). Adaptors, non-speech-motivated movements like touching one's body or manipulating external objects (Litvinenko *et al*., 2018, p. 7), were not included in the analysis. The number of gestures was counted based on the number of gesture phrases, which consist of a preparation phase and a stroke phase (Kendon, 2004). In the case of multiple strokes in one gesture phrase, we calculated each stroke as one gesture.

We then analyzed the distribution of gestures between fluent and disfluent speech. Following Gile (2008), we used sentences as the unit of analysis. Fluent speech was characterized by the absence of any interruptions or disruptions within the sentence boundaries. Thus, gestures accompanied by fluent speech only refer to those gestures that were embedded in a complete fluent sentence, where no disfluencies emerged within the sentence boundaries. We labeled this group as "G + Fluency".

Disfluent speech was defined as a sentence with disfluencies occurring in its boundaries. Gestures parallelling disfluent speech were divided into two groups. In some cases, gestures were embedded in a disfluent sentence but were not parallel with disfluencies. That is to say, gestures and disfluencies co-occurred at different places within a sentence. They were adjacent to each other, as they appeared in the same sentence. We labeled this group as "G + Quasi-disfluency". In other cases, gestures parallelled disfluencies per se. We labeled them as "G + Disfluency". We extended the boundaries of disfluencies a little bit by including the first phoneme or word that immediately preceded or followed the disfluency. This makes sense because the boundaries of gesture and disfluency may not completely overlap with each other, and there could be a short time lag (usually a few milliseconds) between their boundaries. For example, a gesture's onset and preparation phase could be performed within the pause, with the subsequent stroke phase overlapping with the first phoneme after the pause. Gestures and disfluencies partially overlapped with each other. Hence, we also included in this group gestures whose strokes were located on the first phoneme or word that immediately preceded or followed the disfluency.

The "G + Disfluency" group deserves more attention, for these were cases where gestures could be potentially related to the fluctuation of cognitive load. We gave a more detailed account of the temporal relation between gestures and disfluencies for this group. Given that we included the first phoneme or word before and after disfluency into its boundaries, three categories naturally emerged. We used "Pre-disfluency G" to refer to gestures whose stroke co-occurred with the first phoneme or word *before* disfluency. 'Peri-disfluency G" referred to gestures whose stroke was *within* disfluency. "Post-disfluency G" referred to gestures whose stroke was produced with the first phoneme or word *after* disfluency.

In some cases, a gesture may consist of multiple strokes. Some strokes were performed within a disfluency, while other strokes were performed with the following fluent word(s). Strokes that accompanied the fluent word were designated as 'Post-disfluency G', while only those strokes that overlapped with disfluencies per se were classified into the 'Peri-disfluency G' category.

Gesture annotations were conducted following the annotation procedures developed by Litvinenko and colleagues (Litvinenko *et al*., 2018). We calculated gesture rate and disfluency rate, which referred to the total number of gestures and the total number of disfluencies divided by the duration of the interpreting product, respectively. We used minutes as the unit of time. Statistical analyses were conducted in SPSS 26.0.

The coding of gestures and disfluencies was mainly conducted by the first author twice, with a time interval of one month. Intra-coder agreement was 83.9%, indicating a high degree of agreement. After that, the second author randomly coded 5% of the data, following the same procedures. Inter-coder agreement was 78.6%, reflecting a strong level of consistency. The two authors also discussed and resolved inconsistencies.

3. Results

3.1. Disfluency patterns in SI tasks

Altogether, we have identified 3,245 disfluencies in the SI dataset, including 2,262 silent pauses longer than 0.5 seconds (69.7%), 361 filled pauses (11.1%), and 622 false starts (19.2%). Filled pauses are usually comprised of a filler word, like *um* or *uh*, with silent pauses occurring before and/or after the filler word. In our analysis, if the silent part of the filled pause exceeds the threshold of 0.5 seconds, it is coded as a separate silent pause.

Disfluency rate refers to the number of disfluencies per minute (dpm). The mean disfluency rate was 20.03 dpm (*SD* = 3.71 dpm, range 14.05 – 30.97 dpm). Using disfluencies as indicators of cognitive load, we compared the cognitive load between video and audio interpreting conditions. The mean disfluency rate for SI in video condition (*N* = 11) was 19.87 dpm (*SD* = 3.34 dpm, range 14.05 – 25.77 dpm), and the mean disfluency rate for SI in the audio condition (*N* = 11) was 20.18 dpm (*SD* = 4.20 dpm, range 16.29 – 30.97 dpm). Results of the independent sample t-test showed that the disfluency rate in video and audio conditions was not significantly different, $t(20) = -.191$, $p = .850$. The effect size was small ($d = .082$). Although the cognitive load in audio interpreting was slightly higher than that in the video condition, the difference was not statistically significant.

3.2. Gestural behaviors in SI tasks

For gestures, we have identified 317 gestures from SI tasks: 7 deictic gestures, 6 iconic gestures, 57 metaphoric gestures, and 247 beat gestures. Their distribution among fluent and disfluent speech is shown in Table 2:

	Gesture types					
Task types	Deictics	Iconics	Metaphorics	Beats	Total	Percentage
G + Fluency	0	0	0	0	0	0%
G + Quasi-disfluency	5	0	15	90	110	34.7%
G + Disfluency	2	6	42	157	207	65.3%
Total	**7**	**6**	**57**	**247**	**317**	**100%**

Table 2. Number and types of gestures in the SI tasks

It is interesting to notice that all the gestures in SI occurred in disfluent sentences. 65.3% of gestures paralleled disfluencies ('G+ Disfluency'), while the remaining 34.7% was adjacent to disfluencies ('G + Quasi-disfluency'), where disfluencies occurred elsewhere within the sentence and did not overlap with gestures. None of the gestures appeared in a fully fluent sentence ('G + Fluency'). We will further explore the 'G+ Disfluency' category in the next section.

In terms of gesture types, beats were the most frequently used gesture types, which occupied more than three-quarters of the dataset (77.2%), outnumbering all other gesture types. Metaphoric gestures were moderately used with a percentage of 18.1. Deictic (2.2%) and iconic (2.5%) gestures were less frequently used among participants in SI.

Like disfluency rate, gesture rate was calculated as the number of gestures per minute (gpm). The mean gesture rate was 2.25 gpm (*SD* = 2.55 gpm, range 0.16 – 8.87 gpm). T-test showed that the gesture rate in video (*M* = 2.11 gpm, *SD* = 2.49 gpm) and audio (*M* = 2.39 gpm, *SD* = 2.74 gpm) conditions were not significantly different ($t(20) = -.243$, $p = .810$). The effect size was small ($d = .107$). Although participants gestured slightly more during audio interpreting, this difference was not statistically significant, indicating comparable gestural behaviors between video and audio interpreting.

We also tested the correlation between gesture rate and disfluency rate. A Pearson correlation coefficient was computed to measure the linear relationship between gesture rate and disfluency rate. No positive or negative correlations were found between the two variables ($r(20) = .344$, $p = .117$). The effect size was medium ($r^2 = .118$). This indicates that there was no significant linear relationship between gesture and cognitive load in the SI tasks.

3.3. Temporal relation between gesture and disfluency

From the above analysis, we noticed that all the gestures in SI occurred in disfluent sentences, among which 65.3% co-occurred with disfluencies per se. We then focused on this 'G + Disfluency' group, and conducted a detailed analysis of the temporal relation between gesture and disfluency.

There were 207 gestures in this group, including 2 deictic gestures, 6 iconic gestures, 42 metaphoric gestures, and 157 beat gestures. More than half (59.4%, *N* = 123) of these gestures were preceded by disfluencies (the 'Post-disfluency G' category). They overlapped with the first phoneme or word after the disfluency. We have to point out that there was still a partial overlapping between gestures and disfluencies for this category. 79 gestures (38.2%) overlapped with disfluencies per se (the 'Peri-disfluency G' category), and only 5 gestures (2.4%) preceded disfluencies (the 'Pre-disfluency G' category), which means that they overlapped with the first phoneme or word before the disfluency. Like the 'Post-disfluency G' category, these 5 gestures still partially overlapped with the following disfluencies. The distribution of gestures before, within, and after disfluencies is shown in Figure 1.

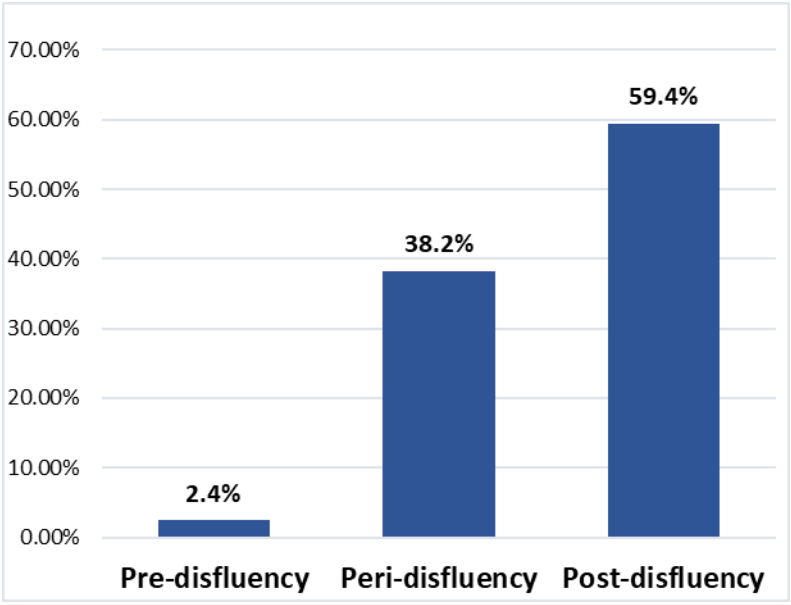


Figure 1. Distribution of gestures before, within, and after disfluencies for the 'G + Disfluency' group (*N* = 207)

We then distinguished whether the disfluency overlapping with gestures was a pause (P) or a false start (FS). We did not differentiate between silent and filled pauses because they had similar functions. 60.9 % of them were pauses and 39.1% were false starts (see Figure 2).

Putting the two figures together, we obtain a more detailed picture, which includes six different sub-categories of the types of gestures and disfluencies, as well as their temporal relations (see Table 3).

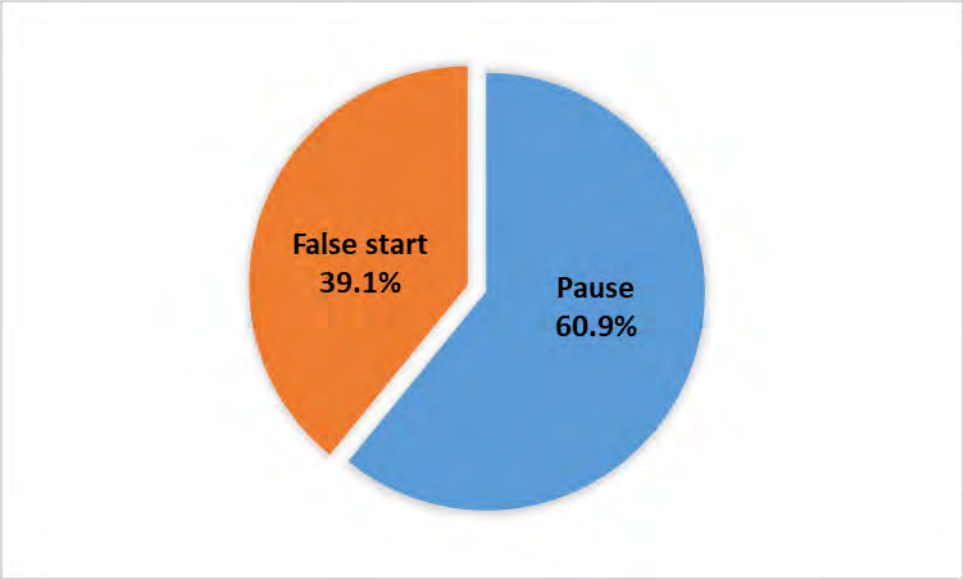


Figure 2. Types of disfluencies overlapping with gestures for the 'G + Disfluency' group (*N* = 207)

	Temporal relations						
	Pre-disfluency G		Peri-disfluency G		Post-disfluency G		
Gesture types	Pre-P	Pre-FS	Peri-P	Peri-FS	Post-P	Post-FS	Total
Deictics	0	0	0	0	0	2	2
Iconics	0	0	0	1	2	3	6
Metaphorics	0	1	10	7	14	10	42
Beats	3	1	42	19	55	37	157
Total	**3**	**2**	**52**	**27**	**71**	**52**	**207**

Table 3. Types of gestures and disfluencies with their temporal relations

Almost all the deictic and iconic gestures occurred *after* disfluency, with only one exception, in which the iconic gesture overlapped with a false start. Only 4 beat gestures and 1 metaphoric gesture occurred *before* disfluency. Metaphoric gestures were evenly distributed between pauses and false starts, where nearly half (*N* = 24) of them were accompanied by pauses, and the other half (*N* = 18) went with false starts. Beat gestures were more closely tied with pauses, with nearly two-thirds (*N* = 100) overlapping with pauses.

4. Discussion

4.1. Relationship between gesture and cognitive load

In addressing the first research question, we will discuss the relationship between gesture and cognitive load in SI from three distinct perspectives.

First, gestures are likely to be produced at moments of processing difficulty. All the gestures in SI were produced in disfluent sentences, among which more gestures were parallel with disfluencies per se (the 'G + Disfluency' group). This result is in line with Stachowiak-Szymczak (2019). In her study, both professional interpreters and trainees produced more gestures when interpreting lists and numbers compared to interpreting narratives. Lists and numbers are problem triggers that require more processing capacity. The result is also partly aligned with that of Adam and Castro (2013) on student interpreters, in which 18.41% of the observed gestures appeared at moments of hesitation in SI, such as pauses or self-correction. This indicates that interpreters are likely to produce gestures when they are experiencing a concrete problem in cognitive processing.

Second, interpreters tend to gesture more in tasks with a higher cognitive load. When comparing different task conditions, disfluency rate in audio interpreting was found to be slightly higher than that in video interpreting. The same trend was found in the comparison of gesture rate between the two conditions. Even though the differences in both cases were not statistically significant, more gesture was used in the condition with a higher cognitive load. The difference of disfluency rate could be explained by the effect of codes and modalities on task interference (Seeber, 2007; 2011). In the audio condition, the input and output processes are both in auditory modality and only verbal processing is involved; while in the video condition, the level of processing underlying visual modality is multimodal and is different from that of auditory modality. Thus, greater interference and more cognitive load would arise in audio interpreting than in video interpreting, because the two processes in audio interpreting have common structures. Furthermore, the fact that more gestures are used in audio interpreting corroborates the facilitative role of gestures (Kita *et al*., 2017). Gesturing while speaking can reduce the cognitive load on working memory (Goldin-Meadow *et al*., 2001).

Third, gestures were embodied, multimodal manifestations of exported load. In our dataset, we found that the use of gestures was parallel with or adjacent to processing difficulty: all the gestures are produced in disfluent sentences. However, the correlation between gesture and cognitive load did not reach a significant level. This could be explained by the notion of 'exported load' (Gile, 2008), which refers to the phenomenon that extra cognitive load may be carried over to downstream segments, leading to cognitive saturation at a later stage. Based on the assumption that speech and gesture are generated from the same cognitive mechanism (Hostetter & Alibali, 2008), exported loads may be reflected in speech as disfluencies or as speech-accompanying gestures.

For the 'G + Disfluency' group, where gestures were in parallel with disfluencies, more than half (59.4%) were produced after disfluency (the 'Post-disfluency G' category). Here, disfluency reflects a moment at which a processing problem arises and the interpreter stops the speech production process to solve the problem. Solving problems requires more effort, leading to an increase in local cognitive load. When the problem is solved and speech production is resumed, the extra load is carried over downstream, which might be manifested in the gestural modality. Thus, for gestures parallel with disfluencies, they tend to appear after disfluency.

For the 'G + Quasi-disfluency' group, gestures could also reflect exported load. These gestures are performed with disfluent sentences, but they are not in parallel with disfluencies per se. Loads from the previous sentence(s) could influence the downstream processing, resulting in disfluencies or gestures.

4.2. Functions of gesture under cognitive load

For the second question, we will discuss the function of each gesture type under cognitive load.

First, in our dataset, the types of gestures produced during silence, i.e. produced with pause, are beats and metaphorics. In the 'Peri-P' sub-category, which refers to gestures parallel with pauses, there are 10 metaphoric gestures and 42 beat gestures. No deictic and iconic gestures are found in this sub-category. Through detailed analysis of these gestures, we find that their usage is in line with the cognitive functions described by other researchers. In the following examples, we followed McNeill's (1992) transcription system to transcribe gestures. For disfluencies, silent pauses are transcribed as a double slash (//) plus their duration in parentheses. When presenting examples, we include both SL and TL as well as a word-for-word English equivalence in italicized form beneath the corresponding character in TL.

The use of metaphoric gesture during silence echoes McNeill's (1985) description of 'silent gesture': such gesture relies on the conceptualization of linguistic units as containers. Example (1) shows the use of a silent metaphoric gesture by participant No. 7 (P07) when interpreting speech A.

Example (1): Silent metaphoric gesture in audio interpreting

SL (A): What you find are three clusters of character strengths.

TL (P07): // (0.7s) 你会发现 **[… // (0.8s)]** **Metaphoric** [三] Beat [种] Beat [不同的] Beat 评价

you will discover *three* *kinds* *different* *evaluations*

In this example, after interpreting the first three words, the participant stopped for 0.8 seconds, and a metaphoric gesture was produced during this silent pause. The form of this gesture is a small finger-lift movement (Cienki, 2021; this volume), where both of her thumbs were raised upward and her right forefinger was stretched forward (see Figure 3I). This is a typical gestural form to present a point when the speaker is seated, with hands on a table in the front, palms facing the speaker (Cienki, 2021, p. 20). Then the participant resumed interpreting and produced three beat gestures, which were miniature tapping movements performed by her right forefinger, along with the subsequent verbal products.

Figure 3. The silent metaphoric gesture in audio interpreting by P07

A plausible explanation was that the participant used a waiting strategy (Seeber, 2011), by which she halted TL production to wait for more input from SL. At the same time, the metaphoric gesture called forth a container in which the meaning of the subsequent segment could be filled. This gesture was not semantically related to the verbal product, but it reflected the problem-solving strategy during silence (Kita *et al*., 2017).

The use of beat gestures during silence is also in line with McNeill's (1985) description: they serve as an attempt to get the speech process going again. Lucero *et a*l. (2014) also found that

beat gestures can facilitate speech production. Example (2) illustrates the use of silent beat gestures by participant No. 2 (P02) when interpreting speech A.

Example (2): Silent beat gestures in video interpreting

> SL (A): So these can be called, in David Brooks' language, the "résumé strengths", because these are the things that get you hired.
>
> TL (P02): // (2.6s) <呃> 我的朋友 / David Brooks **[[…] […] // (7s)]** $_{\text{Beat}}$ **[**嗯**]** $_{\text{Beat}}$ / 他
> *er my friend David Brooks en he*
>
> 认为这是一个长期形成的过程 // (2.6s)
> *thought this is a long formation process*

There was a long silent pause in this segment, which lasted for 7 seconds. During the pause, the participant produced two beat gestures, which were downward movements performed by her right thumb (see Figure 4). Each gesture stroke is marked by a pair of square brackets in transcription.

Figure 4. The silent beat gesture in video interpreting by P02

In the retrospection comment, the participant mentioned that she could not fully understand the meaning of the phrase "résumé strengths" and was searching for a Chinese equivalence, even though she did understand the meaning of the individual word "résumé". She did not come up with an equivalent expression in the TL, and then she resumed interpreting the subsequent segment with a filler word "嗯", which was marked by a third beat gesture.

The two gestures produced during the silent pause also reflected, in an indirect way, the problem-solving strategy. The number of gestures indicates that the participant tried to resume her interpretation twice when the production of interpreting was halted. The participant's retrospection echoes the finding of Adam and Castro (2013). When beat gestures were produced at moments of hesitation, the interpreter was either having a comprehension problem or engaging in word search.

The third gesture produced with the filler word indexed the end of silence and the beginning of the following interpretation. These three gestures are not semantically related to verbal products, but they could serve an emphasizing function in that they implicitly contrast the absence of a word with its desired presence (McNeill, 1985, p. 359). In this sense, silent beat gestures could have a meta-cognitive function, namely monitoring.

Second, deictic and iconic gestures are not produced before the semantically related linguistic items but rather parallel with them. In McNeill's (1985) illustration of gestural anticipation, a gesture could be produced prior to the production of its corresponding linguistic item. However, in our dataset, we found that deictic and iconic gestures were produced with the corresponding word itself. The anticipation of gestures is not obvious. Example (3) shows the use of iconic gestures by participant No. 1 (P01) when interpreting speech B.

Example (3): Iconic gesture in video interpreting

SL (B): When the work came back, I calculated grades.

TL (P01): // (0.6s) 之后 / **[**收回来**]** $_{\text{Iconic}}$ 这些任务之后呢我开始去算他们的分数 // (1.2s)

then collect back these tasks after I began to calculate their grades

Figure 5. The iconic gesture in video interpreting by P01

When producing the phrase "收回来", the participant produced an iconic gesture, which was a backward movement toward her own body, mimicking the action of "collect" (see figure 5). This gesture was semantically redundant, for it expressed the same linguistic meaning with speech. In terms of temporal relations, the gesture is produced simultaneously with its linguistic counterpart, not preceding it.

Deictic and iconic gestures are semantically related to concurrent speech (Arbona *et al*., 2023). There were only 2.2% (*N* = 7) deictic gestures and 1.9% (*N* = 6) iconic gestures in the dataset, including the one in Example (3), and all were produced with linguistic items. However, in similar works, simultaneous interpreters produced more of such semantically related gestures (Martín de León & Fernández Santana, 2021; Zagar Galvão, 2020). This could be influenced by factors such as the content of the speech, the speaker's style, and cultural differences.

The lack of anticipation for semantically related gestures could be explained by the task peculiarity of SI. Unlike in spontaneous speech production, the meaning of the interpreting product comes not from the interpreter, but from the speaker. The interpreter has to receive, and perhaps wait for, the input from the speaker, while producing the output in another language. This mental process is different from that of gestural anticipation in spontaneous speech. Given that the cognitive demands in SI are high, there is not enough working memory capacity to store the pre-activated representations for gestures. Such representations are only activated when producing the verbal product and simultaneously produced in gestures. The production of deictic and iconic gestures could be affected by cognitive load.

5. Conclusion

This study explored the relationship between gesture and cognitive load in simultaneous interpreting. Our findings are as follows. First, gestures in SI are likely to be produced with processing difficulty, especially when interpreters are experiencing a concrete problem. Even if the correlation between gesture and cognitive load is not statistically significant, interpreters tend to gesture more in tasks with a higher cognitive load. This corroborates the facilitative role of gestures in cognitive processing. Based on the assumption that speech and gesture are generated from the same cognitive mechanism, gestures could be an embodied, multimodal manifestation of 'exported load'. The temporal relations between gesture and disfluency give support to this claim. Processing difficulty could result in a speech breakdown and a parallel gesture is produced. The strokes of this gesture tend to follow the speech disfluency.

Additionally, processing problems could also influence downstream segments, leading to both disfluencies and gestures. However, in such cases, gesture and disfluency co-occur within a sentence boundary, but they do not overlap. Both of the two modalities could be reflections of exported loads. Second, beat and metaphoric gestures are connected with processing difficulty. Silent beat gestures could reflect the interpreter's monitoring of the interpreting process, which means that beats might have a meta-cognitive function. On the other hand, silent metaphoric gestures may reflect the problem-solving strategy. When producing metaphoric gestures, interpreters stopped TL production and called forth a container for the subsequent speech segments to fill in. When speech production comes to a halt, cognitive processing does not stop. Silent gestures are embodied manifestations of such processes. Third, deictic and iconic gestures could convey concurrent semantic meaning, but such gestures are less frequently used by the interpreters of the study. Due to the task peculiarity of SI, there is a lack of anticipation for such gestures. Their production could be affected by cognitive load. Semantically related gestures provide a multimodal perspective for studying the cognitive processes of interpreting.

This study contributes to the understanding of SI as an embodied, multimodal cognitive activity. The findings have indicated several lines of research in the future. There is a difference between the number of gestures produced by interpreters in this study and by interpreters from other countries. A comparative study could unveil the cultural differences behind gestural styles. Future research could also make comparisons between gestures in SI and CI to further explore the relationship between gesture and cognitive load. One limitation of this study is the number and the level of professional competence of participants. In the future, we will recruit more participants as well as professional interpreters. A comparison between professional and student interpreters will also shed light on the effect of interpreting competence on gestural behaviors.

6. Acknowledgement

This research is supported by "the Fundamental Research Funds for the Central Universities" in UIBE (23QD12). We would like to thank professor Alan Cienki, Dr. Celia Martín de León, Dr. Sílvia Gabarró-López and the two anonymous reviewers for their valuable comments and suggestions on earlier drafts of this paper. The responsibility for the content and any remaining errors remains exclusively with the authors.

7. References

Adam, C., & Castro, G. (2013). Schlaggesten beim Simultandolmetschen – Auftreten und Funktionen, *Lebende Sprachen, 58*(1), 71–82. https://doi.org/10.1515/les-2013-0004

Arbona, E., Seeber, K., & Gullberg, M. (2023). Semantically related gestures facilitate language comprehension during simultaneous interpreting. *Bilingualism: Language and Cognition*, *26*(2), 425–439. https://doi.org/10.1017/S136672892200058X

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bóna, J., & Bakti, M. (2020). The effect of cognitive load on temporal and disfluency patterns of speech: Evidence from consecutive interpreting and sight translation. *Target*, *32*(3), 482–506. https://doi.org/10.1075/target.19041.bon

Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives*, *25*(4), 640–657. https://doi.org/10.1080/0907676X.2016.1278026

Christoffels, I., & De Groot, A. (2005). Simultaneous interpreting: A cognitive perspective. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 454–479). Oxford University Press.

Cienki, A. (2021). From the finger lift to the palm-up open hand when presenting a point: A methodological exploration of forms and functions. *Languages and Modalities*, *1*, 17–30. https://doi.org/10.3897/lamo.1.68914

Cienki, A., & Iriskhanova, O. K. (2020). Patterns of multimodal behavior under cognitive load: An analysis of simultaneous interpretation from L2 to L1. *Voprosy Kognitivnoy Lingvistiki*, *1*, 5–11. https://doi.org/10.20916/1812-3228-2020-1-5-11

Gile, D. (1999). Testing the Effort Models' tightrope hypothesis in simultaneous interpreting – A contribution. *Hermes*, *23*, 153–172. https://doi.org/10.7146/hjlcb.v12i23.25553

Gile, D. (2008). Local cognitive load in simultaneous interpreting and its implications for empirical research. *Forum*, *6*(2), 59–77. https://doi.org/10.1075/forum.6.2.04gil

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychological Science*, *12*(6), 516–522. https://doi.org/10.1111/1467-9280.00395

Gumul, E. (2021). Explicitation and cognitive load in simultaneous interpreting: Product- and process-oriented analysis of trainee interpreters' outputs. *Interpreting*, *23*(1), 45–75. https://doi.org/10.1075/intp.00051.gum

Han, C., & An, K. (2020). Using unfilled pauses to measure (dis)fluency in English-Chinese consecutive interpreting: In search of an optimal pause threshold(s), *Perspectives*, *29*(2), 1–17. https://doi.org/10.1080/0907676X.2020.1852293

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514. https://doi.org/10.3758/PBR.15.3.495

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245–266. https://doi.org/10.1037/rev0000059

Litvinenko, A. O., Kibrik, A. A., Fedorova, O. V., & Nikolaeva, J. V. (2018). Annotating hand movements in multichannel discourse: Gestures, adaptors and manual postures. *The Russian Journal of Cognitive Science*, *5*(2), 4–17.

Lucero, C., Zaharchuk, H., & Casasanto, D. (2014). Beat gestures facilitate speech production. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the cognitive science society* (pp. 898–903). Curran Associates.

Martín de León, C., & Fernández Santana, A. (2021). Embodied cognition in the booth: Referential and pragmatic gestures in simultaneous interpreting. *Cognitive Linguistic Studies*, *8*(2), 277–306. https://doi.org/10.1075/cogls.00079.mar

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350–371. https://doi.org/10.1037/0033-295X.92.3.350

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.

Muñoz Martín, R. (Ed.). (2016). *Reembedding translation process research*. John Benjamins.

Seeber, K. G. (2007). Thinking outside the cube: Modeling language processing tasks in a multiple resource paradigm. *Interspeech 2007, 8th Annual Conference of the International Speech Communication Association* (pp. 1382–1385). https://doi.org/10.21437/Interspeech.2007-21

Seeber, K. G. (2011). Cognitive load in simultaneous interpreting: Existing theories - new models. *Interpreting*, *13*(2), 176–204. https://doi.org/10.1075/intp.13.2.02see

Seeber, K. G. (2013). Cognitive load in simultaneous interpreting: Measures and methods. *Target*, *25*(1), 18–32. https://doi.org/10.1075/target.25.1.03see

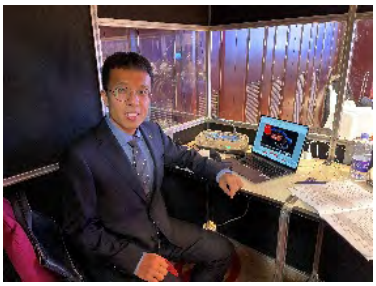
Seeber, K. & Kerzel, D. (2012). Cognitive load in simultaneous interpreting: Model meets data. *International Journal of Bilingualism*, *16* (2), 228–242. https://doi.org/10.1177/1367006911402982

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. *Proceedings of the 6th international conference on language resources and evaluation (LREC 2008)* (pp. 816–820).

Stachowiak-Szymczak, K. (2019). *Eye movements and gestures in simultaneous and consecutive interpreting*. Springer.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636. https://doi.org/10.3758/BF03196322

Zagar Galvão, E. (2020). Gesture functions and gestural style in simultaneous interpreting. In H. Salaets & G. Brône (Eds.), *Linking up with video: Perspectives on interpreting practice and research* (pp. 151–179). John Benjamins. https://doi.org/10.1075/btl.149.07gal



Yuetao Ren
University of International Business and Economics
Huixin Dongjie, 10
100029 Beijing
China

renyuetao@gmail.com

Biography: Yuetao Ren is a lecturer in Translation and Interpreting at the University of International Business and Economics (UIBE), China. He received his PhD degree in Interpreting Studies from Renmin University of China. His research interests include interpreting studies, gesture studies, and multimodality. He also works as the coordinator for the conference interpreting program at UIBE and as a freelance conference interpreter.

Jianhua Wang
Renmin University of China
Zhongguancun Street, 59
100872 Beijing
China
wjhsfl@ruc.edu.cn

Biography: Jianhua Wang is a professor of Translation and Interpreting at Renmin University of China (RUC). He holds a PhD degree in Cognitive Psychology. His current research interests include the cognitive process of translation and interpreting, translation and communication, multimodal translation, discourse studies, as well as global and area studies. He also serves as the associate dean of the School of Foreign Languages at RUC.

Parallèles 37(1) April 2025

Modeling gestural alignment in spoken simultaneous interpreting: The role of gesture types

Inés Olza

Institute for Culture and Society. University of Navarra

Abstract

This article explores gestural alignment in spoken simultaneous interpreting, analyzing whether and how the interpreters under scrutiny align with the gestural behavior of a visible speaker-source, and which gesture types by the speaker-source more often prompt a gesturally aligned response by the interpreters. The paper offers a mixed-methods analysis of a set of multimodal data collected under (quasi-)experimental conditions in a real court interpreting setting during spoken training exercises performed by two novice interpreters. This study relies on the findings of a previous exploratory approach to the same dataset (Olza, 2024), where different degrees of gestural alignment were found and defined. In this study, the variable *gesture type* is used to systematically examine a new sub-sample of the same data and to compare the performance of the two novice interpreters. Results show that iconic gestures elicit higher degrees of alignment by both interpreters. The findings are not conclusive, though, when relating the (non-)representational nature of gestures by the speaker-source, nor their (non-) semantic value, to the degree of replication of such gestures by the two interpreters. Future research will rely on broader datasets obtained from more experienced interpreters engaged in tasks that more accurately reflect their actual practice.

Keywords

Gesture, alignment, spoken simultaneous interpreting, multimodal data, gesture types

 DOI:10.17462/para.2025.01.05

1. Introduction

1.1. Gesture in spoken simultaneous interpreting: previous studies

This paper aims to contribute to the growing body of research on gesture in spoken and signed-to-spoken simultaneous and consecutive interpreting, where a 'multimodal turn' in training, practice and analytical approaches is proposed, thanks mainly to the pervasive presence and use of video in professional and academic settings (Salaets & Brône, 2020). Indeed, the possibility, for the interpreters, of fully accessing and watching the speaker-source's performance (including non-verbal behavior) and, for the researchers, of recording and scrutinizing both the speaker-source's and the interpreter's multimodal behavior, makes it possible to conduct in-depth systematic analyses of gesture in interpreting tasks (Chwalczuk, 2021; Stachowiak-Szymczak, 2019; Zagar Galvão, 2015), and to relate the gestural behavior of the interpreters to that of the speaker-source (Olza, 2024; Zagar Galvão, 2013).

Specifically, empirical research on gesture within the subfield of spoken simultaneous interpreting[1] has focused thus far on three complementary strands.

(a) Studies aiming to define the presence and role of gesture within the overall performance of professional interpreters and/or trainees, mainly to determine its role in processing the cognitive load demanded by interpreting tasks in experimental (Stachowiak-Szymczak, 2019, chapters 5 and 6) and naturalistic settings (Chwalczuk, 2021; Fernández Santana & Martín de León, 2022; Iriskhanova et al., 2023; Martín de León & Fernández Santana, 2021; Zagar Galvão, 2015, 2020), with elicited or real interpreting tasks.

(b) Research integrating the advances in multimodal (interaction) studies into the analysis of the interpreters' performance, including fine-grained analyses of the deployment of certain gesture types in their behavior (e.g. gaze and beat gestures, in Stachowiak-Szymczak, 2019; iconic gestures, in Fernández Santana & Martín de León, 2022; metaphoric gestures, in Leonteva et al., 2023); the role of gesture in managing interpreting disfluencies (Cienki, 2024); or the complex interactional dynamics shaping gesture and language in interpreting and interpreter-mediated contexts (Krystallidou, 2020), among other perspectives.

(c) Incipient research on the degree of gestural convergence between the speaker-source and the interpreter, with several qualitative and mixed-methods analyses providing preliminary empirical evidence of how interpreters often reproduce in their own discourse the gestures they observe in the speaker-source (Chwalczuk, 2021; Janzen et al., this special issue; Leonteva et al., 2023; Zagar Galvão, 2013;). Within this strand of research, especially iconic and metaphoric gestures were analyzed across the speaker-source and the interpreters' performance (Chwalczuk, 2021, section 4.1.1; Leonteva et al., 2023).

Within this background, our study aims to add to the understanding of the cognitive mechanisms that regulate spoken simultaneous interpreting, with a focus on how the gestural convergence exhibited by the interpreters with regard to the speaker-source may be described and discussed from the tenets of both alignment theories (section 1.2) and gesture studies (namely, those leading to the definition and characterization of gesture types; section 2.3.1).

[1] A recent overview of the research on gesture in other interpreting modalities (signed-to-spoken interpreting) and types (consecutive and distance interpreting), as well as in spoken simultaneous interpreting, was offered in the panel 'Gesture in spoken and signed-to-spoken language interpreting', convened by Sílvia Gabarró-López and Alan Cienki at the 18th International Pragmatics Conference (2023 IPrA Conference, Université libre de Bruxelles, July 2023), where around 20 scholars in pragmatics, cognitive and applied linguistics discussed the latest advancements in gesture analysis in interpreting. A first approach to the dataset of this study was presented at this panel. I would like to thank the convenors and participants for their insightful comments and suggestions.

1.2. Cognitive and gestural alignment: previous approaches and key definitions

In this paper, gesture is viewed as both a reflection and a shaper of human thought, serving as a material anchor to explore the embodied mental mechanisms that underlie language use (Cienki, 2022; McNeill, 1992, 2005). One of these large-scale cognitive operations that unfold across diverse kinds of human behavior, including language, is alignment. From an external perspective that observes a given subject's behavior, alignment encompasses the dynamics of convergence and divergence of his/her actions relative to others. These behavioral 'movements' materialize in changes and adaptations (accommodation) of his/her communicative behavior at different levels (verbal, paraverbal, non-verbal) (Giles & Ogay, 2007, p. 295). In other words, while engaging in linguistic interaction, speakers monitor their own behavior and that of their interlocutors, and consequently —even 'strategically'—approach (align with) or distance (misalign) their behavior relative to that of others. This adaptation may occur at a local level, through alignment in specific linguistic and gestural choices, or unfold in a sustained and progressive manner throughout an entire conversational exchange (Fusaroli & Tylén, 2016).

Such a definition of alignment was first proposed within the framework of Communication Accommodation Theory (CAT), rooted in social psychology and sociolinguistics (Giles et al., 1991; Giles & Ogay, 2007). It was later revisited and expanded by cognitive and behavioral studies, which have shown that (mis)alignment regulates not only face-to-face interaction and communication but also all kinds of human behavior involving cooperation between individuals. This includes joint actions ranging from physical manipulation of objects (e.g., cooking together) to symbolic tasks (e.g., playing together). Within this cognitive and behavioral framework, alignment has been defined from two complementary perspectives.

First, it has been described and explained as a material manifestation of the wider *priming* principle that regulates human interaction, taken as "an automatic, bidirectional process operating in parallel on several different levels of representation" (Healey, 2004, p. 201), through which the interacting individuals —the interlocutors, in the case of communication— couple their respective situational models, that is, their mental representations of the situation and/or issues under discussion (Pickering & Garrod, 2004, sections 2.1-2.3). As a result, interlocutors not only cooperate during interaction but also align through a form of 'mimetic' behavior, where they converge by 'imitating' each other's actions. Going beyond the logics of stimuli-response underlying the described priming views (Doyle & Franck, 2016; Krauss & Pardo, 2004), the second big approach to alignment proposes to analyze it under the scope of *grounding* and interpersonal synergy (Fusaroli & Tylén, 2016), as a form of synchronized activity which is negotiated in a relational way, with wider room for the joint attention and cooperative action (Eilan et al., 2005; Goodwin, 2018) that characterize any form of human communicative exchanges. Here, alignment strongly relies on the common goals and common ground, and the communicative dynamics established between the interlocutors in concrete, genre-based, and situated interaction, in a similar way to how conversational analytic approaches describe them (Riordan et al., 2014; Stivers, 2008; Stivers et al., 2011). All in all, both approaches stress that the participants engaged in communicative interactions tend to coordinate and converge, i.e., align in their behavior, exhibiting various degrees of mutual 'mimesis' at all linguistic levels (phonetic, lexical, syntactic, semantic), including the gestural one, which has remained largely unaddressed in alignment and accommodation studies until recently (Bergmann & Kopp, 2012; Kimbara, 2006; Kopp & Bergmann, 2013; Rasenberg et al., 2020), most of these recent studies focusing on data elicited and collected in laboratory settings. While contributing to bridge the gap in the study of gestural alignment in interaction, this article aims to address it through the analysis of ecologically valid data of live exercises by novice interpreters conducted in a real, naturalistic setting.

Furthermore, this study relies on the existing body of research on the 'interactive' nature of simultaneous interpreting, that is, the special cognitive and behavioral relationship between the speaker-source, the interpreter, and the recipient of the interpreter's performance, which is claimed to affect the gestural behavior of the interpreter (Chwalczuk, 2021, section 4.1.1; Janzen et al., this special issue; Leonteva et al., 2023). In line with these studies, we assume that interpreters align not only with regard to the speaker-source but also towards the recipients of their performance. In some cases, this could explain why they do not fully align with the speaker-source's gestural behavior, as other types of gestures might be better understood by their audience, as shown, for instance, by Janzen et al. (this special issue). Also, another obvious fact should be noted: even if interpreters seek to 'maximally align' with the speaker-source and his/her behavior and frame of understanding, they do *not* actually *interact* with him/her, at least in the sense that prevails in accommodation and alignment studies, where 'regular' communicative exchanges, that is, those with a dynamic exchange of speaker-listener roles between the interlocutors are examined. This would also explain why interpreters often reproduce 'self-adapted' versions of the gestures carried out by the speaker-source; for instance, simplified gestures that match better with time and cognitive constraints (Leonteva et al., 2023), or gestures that blend their own perspective with that of the speaker-source (Janzen et al., this special issue). Accordingly, although this study mainly focuses on the role of gesture types in modeling gestural alignment in simultaneous interpreting, these distinctive features of the interpreting tasks will also be taken into account in the discussion of results (see section 4).

2. Study design

This study relies on the findings of a previous exploratory approach (Olza, 2024) to the same dataset that is examined here (see section 2.1). In this first study, we conducted a mixed-methods analysis of spoken interpreting data audiovisually recorded in a natural professional setting (courtroom). The study quantified in a basic descriptive way, qualified and compared the degree of gestural alignment towards the same speaker-source exhibited by two distinct novice interpreters, who were recorded while working at the same time in the mentioned setting. The results of this previous research included a taxonomy of the different degrees of gestural alignment found in the data (see also section 2.3), with a good number of instances where the observed interpreters actually mimicked the speaker-source's gestures in type, form and function. Finally, the data analyzed in Olza (2024) were also categorized in an exploratory manner according to several basic gesture types (iconic and metaphoric gestures; discourse-structuring gestures; gestures for modality and stance), which allowed to formulate hypotheses on the higher or lower tendency of certain gesture types to be replicated by the interpreters. In the present paper, these hypotheses on the influence of gesture types on gestural alignment are retaken, expanded and tested in a more granular and systematic way. In sections 2.1 and 2.2, the main features of the study design are presented against the backdrop of the analysis conducted in Olza (2024), so as to explain how the present study advances the understanding of the gesture types that more often prompt an aligned response by the simultaneous interpreters in our data.

2.1. Data

The multimodal data examined here and in Olza (2024) were obtained at real training sessions for novice legal interpreters organized by the interpretation directorate of an official

international court[2]. The interpreters were postgraduate fellows immersed in a specialized training program aimed at integrating them into legal interpreting booths in international institutions. A complete approx. 30 min training session was recorded. This consisted of a live interpreting exercise carried out in a real medium-sized courtroom, where the speaker-source (male) sat at the main orator's position (central front) and the trainees (four subjects) occupied separate booths in both sides of the room. In addition to the trainers (experienced interpreters) of the four novice interpreters, who were sitting next to them in their respective booths, there was no external audience in the room. Due to equipment limitations, only two trainees were recorded.

It should be noted that the speaker-source delivered a speech in Spanish on non-legal issues related to the history of technology. In fact, although they were held by and for court interpreters, the main speech in the training sessions at this particular institution did not necessarily deal with legal issues —it could describe or explain any kind of issue, as happens in our data, where the speaker-source exposed the history of the Thermomix and the dishwasher. This type of non-specialized exercise was usually conducted in the first stages of the training program. In our data, two novice interpreters were recorded: Interpreter 1 (female) worked from Spanish into spoken English; and Interpreter 2 (female) worked from Spanish into spoken French. Visual access to the speaker was similar for both interpreters, as shown below in Figure 1. Three cameras recorded the training and were situated respectively at the right of the main speaker, and directed at both recorded interpreters. The cameras did not interfere with nor block the activity and visual access of the participants. In fact, it is important to note that, during the exercise, the interpreters directed their gaze towards the speaker-source most of the time. The only times they did not look at him were when they appeared to be writing down dates, numbers, and proper names on the papers in front of them. This direction of their gaze thus reinforces the hypothesis that their gestural behavior was aligned with that of the speaker, and was not a result of chance, for example.

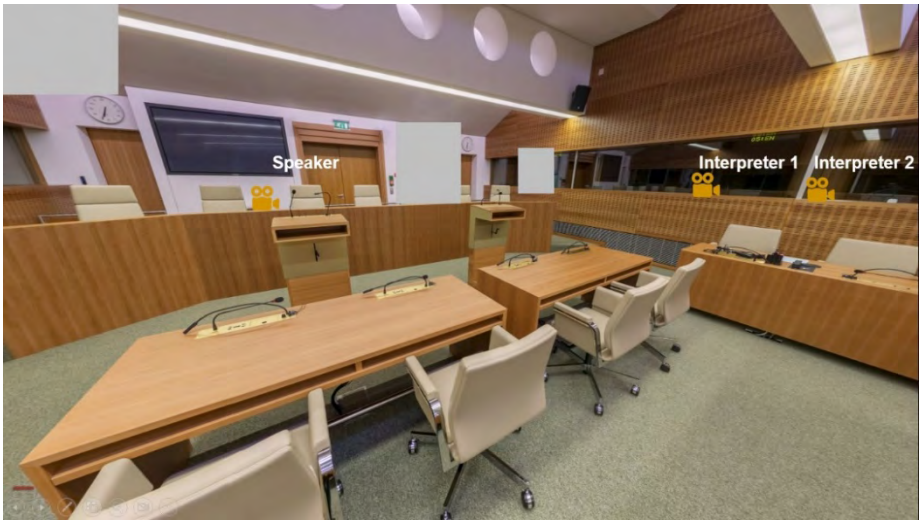


Figure 1. Recording setting (real courtroom): speaker-source, interpreters, and video-cameras

[2] The name and coordinates of the institution are not facilitated due to EU regulation on data protection. Before the sessions, the researcher in charge of the study presented it to the participants, who were able to ask any questions they had before signing the corresponding informed consent form. Previously, the study had received approval from the Research Ethics Committee of the University of Navarra (approval certificate nr. 2017.021).

2.2. Research questions and hypotheses

The analysis was guided by the following research questions and hypotheses, which emerge from the state-of-art described in section 1 and seek to improve and expand the results obtained in our previous study (Olza, 2024).

- **Research question 1.** Do the different gesture types by the speaker-source prompt diverse degrees of alignment by the interpreters in our data?

  Hypothesis 1. *Gestures that do not relate to the speech content (beats, self-adaptors) prompt lower degrees of gestural alignment by the interpreters.*

  The hypothesis is supported by the view of interpreting tasks as discourse (speech) oriented activities, where a common ground of understanding is negotiated with the speaker-source and the recipients of such tasks. Therefore, it seems reasonable to think that gestures not relating to the speech content and structure, and more dependent on the individual style of the speaker-source, will be less often replicated by the interpreters.

  Hypothesis 2. *Within the realm of gestures relating to the representational (referential) or pragmatic (metadiscursive) meaning of speech, iconic gestures and discourse structuring gestures prompt higher degrees of gestural alignment.*

  This hypothesis relies on the results of recent studies that have preliminarily suggested that iconic gestures are very often mirrored by interpreters (Chwalczuk, 2021, section 4.1.1; Olza, 2024), followed by gestures with discourse structuring functions (deictic gestures pointing to discourse referents) (Olza, 2024). In the latter (Olza, 2024), the initial hypothesis was that only representational gestures (iconic and metaphoric) would prompt higher degrees of alignment by the interpreters, as they relate to the referential and conceptual content of discourse. However, in this previous study, discourse structuring gestures were more often replicated than metaphoric gestures. Therefore, relying on a different sub-sample within the same dataset, the present study aims to test and, possibly, replicate the results obtained in Olza (2024).

  Hypothesis 3. *Compared to other types of representational gestures, metaphoric gestures prompt lower degrees of gestural alignment.*

  The hypothesis accords with the unexpected results of our previous study (Olza, 2024), which may replicate here, and those by Leonteva et al. (2023), who show that, due to the cognitive load and time pressure of the tasks, interpreters tend to lower the cognitive complexity (e.g. mental imagery) of their gestural behavior as a response to metaphoric gestures by the speaker-source. In other words, they tend to 'simplify' their gestural output compared to that of the speaker.

- **Research question 2.** Does gestural alignment rely on individuals? Or, on the contrary, does it work similarly in both interpreters who were observed?

  Hypothesis 4. *The degree of gestural alignment exhibited by both interpreters is different due to personal styles and/or differences in fluency and performance quality.*

  The hypothesis emerges from the results of analyzing a different data subsample in Olza (2024), where Interpreter 1 and Interpreter 2 showed different degrees of gestural alignment towards the same speaker-source. The present study seeks to test these findings in a different subsample of the same dataset.

In summary, the research questions and hypotheses presented here aim, on the one hand, to replicate the main results obtained in Olza (2024), particularly to confirm or refute the differences observed in the interpreters' performance, and to test once again the notable

frequency with which iconic gestures were replicated by the novice interpreters. On the other hand, the present study makes novel progress in two directions: in a better systematization of the types of gestures analyzed, with the introduction of gestures unrelated to speech content (beats, which are numerous in the data; and self-adaptors); and in the introduction of the variable 'gesture related/unrelated to speech content' in the study design and discussion of results.

2.3. Methods

The three video recordings (Speaker-source, Interpreter 1, Interpreter 2) were analyzed and tagged separately using the annotation tool ELAN-6.5[3]. The analysis was run by a single coder (the author of this paper) according to the following steps.

2.3.1. Analyzing the speaker-source's gestural behavior: Sample and gesture types

Four 1-minute excerpts of the speaker-source's behavior were analyzed and later on used as a baseline for the comparative analysis of the performance of Interpreters 1 and 2. The excerpts were chosen randomly using an open-source aleatory choice generator[4], resulting in minutes 15:00-16:00, 17:00-18:00, 19:00-20:00 and 21:00-22:00[5]. The selected excerpts were qualitatively analyzed in ELAN. First, the presence of any gesture relevant to the speech content was annotated. The gestures relating to the speech content were temporally delimited and annotated using the tags [gesture type], [body part(s) involved], and [speech sequence going along with gesture].

The coding of gesture types becomes especially relevant for this study, as it revises and expands the gesture types coded in Olza (2024), where a first approach to the influence of gesture types by the speaker-source on the gestural performance of interpreters was offered. In Olza (2024), only representational and pragmatic gestures were distinguished and coded, relying on the following well-established categories (McNeill, 1992; Kendon, 2004): *iconic* —gestures exhibiting a close formal relationship to what is semantically conveyed in speech; *metaphoric* —gestures depicting a figurative image of an abstract concept; *discourse and information structure* —gestures pointing to the discourse referents/topics and/or relating to discourse structuring information; *modality and stance* —gestures for intensification or mitigation of the expressed content; and *gestures for negation*. Beats and self-adaptors (Ekman & Friesen, 1972) were excluded in this exploratory approach, as they do not relate to what is signaled or represented by speech. In this first study (Olza, 2024, section 4), the novice interpreters were observed to align more often with iconic gestures and gestures related to modality and stance (mainly, gestures for intensification of semantic properties conveyed in the speech sequence), with a notable degree of alignment observed also for gestures having discourse structuring functions (e.g. signaling enumerations). Metaphoric gestures by the speaker-source were the type less often replicated by the interpreters.

The results obtained in Olza (2024) had, however, several limitations. First, the complete gestural behavior of the speaker-source was not analyzed, as only gestures related to the speech content were tackled. As pointed out before, beat gestures and adaptors were not coded. Second, the category of gestures for modality and stance turned out to be more problematic than expected, as it involved a higher level of interpretation compared to the other categories, with formally diverse gestures to which the function of expressing the speaker's attitude was

[3] https://archive.mpi.nl/tla/elan (accessed on 15 August 2023).

[4] https://randomchoicegenerator.com/ (accessed on 23 February 2024).

[5] The four excerpts randomly selected and analyzed in Olza (2024) were different (2:00-3:00; 10:00-11:00; 20:00-21:00; and 27:00-28:00).

attributed. For example, the sweeping away gesture (palm down hands moving away in front of the speaker, metaphorically clearing his/her personal space; Bressem & Müller, 2014) was frequently interpreted as a modal gesture for intensification (meaning 'totally'), which resulted in extracting it from the category to which it originally belonged (metaphoric gesture).

To overcome these limitations, the present study analyzes *all* the gestures by the speaker in the selected excerpts, encompassing gestures related to the referential and pragmatic meaning of the speech component (Kendon, 2004, chapter 10), as well as those not relating to the speech content (beats and self-adaptors). Another advance with respect to Olza (2024) has to do with the final set of gesture types analyzed and annotated in the present study, whose definition relies on formal-functional criteria that seek to minimize interpretative biases. Thus, the behavior of the speaker-source was categorized according to the following gesture types (Ekman & Friesen, 1972; McNeill, 1992; Kendon, 2004).

Beats	Gestures that go along with the rhythmical pulsation of speech (i.e., prosody).
Deictic (discourse structure)	Gestures signaling or pointing to the discourse referents/topics.
Head shakes (negation)	Lateral head movements (prototypical gesture for negation).
Iconic	Gestures exhibiting a close formal relationship to what is semantically conveyed in speech.
Metaphoric	Gestures depicting a figurative image of an abstract concept.
Self-adaptors	Non-signaling gestures where one part of the body is applied to another part of the body, such as scratching one's head and face.

Table 1. Analysis of gestural behavior: gesture types

A basic descriptive quantification of the total number of gestures and gesture types that were identified within the speaker-source's excerpts is displayed below in Table 2. In total, 118 gestural units were identified and classified.

Speaker-source		
Gesture type	**Hits**	**Rate**
Deictic (discourse structure)	40	34%
Beat	24	20.3%
Iconic	24	20.3%
Metaphoric	24	20.3%
Adaptor	5	4.2%
Head shake (negation)	1	0.9%
Total	**118**	**100%**

Table 2. Speaker-source: total number of gestures and gesture types

2.3.2. Tracking the gestural response of the interpreters

In a second phase, we analyzed and annotated the interpreters' performance in the excerpts where they interpreted the speech uttered by the speaker-source in the minutes analyzed in the first phase (section 2.3.1). For Interpreter 1, minutes 19:09-20:09, 21:10-22:12, 23:10-24:10, and 25:10-26:10 were analyzed; for Interpreter 2, minutes 17:02-18:02, 19:04-20:05, 21:00-22:00, and 23:00-24:07 were examined. As mentioned above, the speaker-source's behavior was taken as a baseline to analyze the interpreters' performance. Therefore, the sequences

where the speaker-source gestured —sequence = verbal cue and relevant gesture going along with it— were contrasted with the corresponding interpretation by both interpreters. Their behavior in the corresponding sequences was qualitatively analyzed and annotated using the following tags.

Speech interpreted—same type of gesture	While interpreting the verbal-gestural sequence of the speaker-source, the interpreters perform the same kind of gesture as the speaker.
Speech interpreted—different type of gesture	While interpreting the verbal-gestural sequence of the speaker-source, the interpreters perform a gesture of a different type than that of the speaker-source.
Speech interpreted—no gesture	The interpreters translate the verbal sequence that is accompanied by a gesture in the speaker-source's performance, but they do not gesture themselves.
Speech not interpreted	The concrete speech sequence of the speaker-source is not interpreted by the professional, due to disfluencies or constraints in time and expertise.

Table 3. Analysis of interpreters' behavior: tags to define their degree of gestural alignment

These tags sought to track the overall degree of gestural alignment exhibited by both interpreters in response to the verbal and gestural cues observed in the speaker-source. Although it would indeed be relevant to incorporate them into future studies, further details such as the verbal behavior of the interpreters—the actual words used to interpret the verbal sequence under scope—or a thorough formal description of each gesture were not systematically coded, as the main aim of our analysis was to offer a comprehensive comparative approach of the degree of convergence in the gestural behavior of the speaker-source and the interpreters. All in all, as shown in Tables 2 (above, section 2.3.1) and 5 (below, section 3), a total of 270 gestural units were identified and classified in the study: 118 for the speaker-source, 72 for Interpreter 1, and 80 for Interpreter 2.

Returning to the coding methods, cases where the interpreter would perform a similar gesture as the speaker while not interpreting his discourse were not clearly found in our data. Instances where speech was not interpreted were due to disfluencies that did not allow the interpreter(s) to tackle the sequence at all. Consequently, they would simply skip to the next discourse chunk.

The tags included in Table 3 allowed to define a gradual typology of (non-)aligned behavior by the novice interpreters with respect to that of the speaker-source (see Figure 2). Examples of the different degrees of gestural alignment within the proposed continuum are included below (see Figures 3-5). In these examples, and in the remainder of the article, cases where the interpreters did not interpret the verbal sequence produced by the speaker will not be considered.

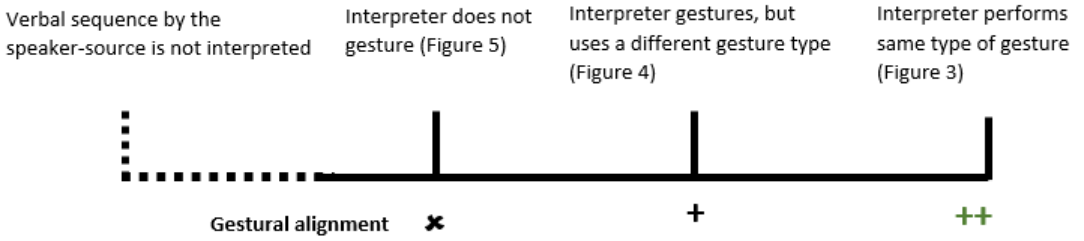


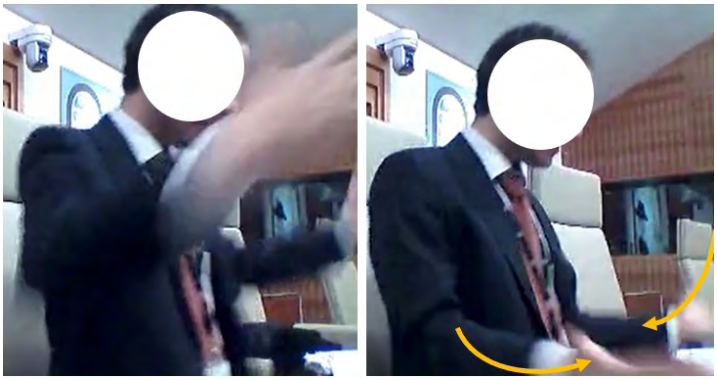
Figure 2. The 'gestural alignment continuum' (adapted from Olza, 2024)

Highest gestural alignment (++): Interpreter uses the same gesture type as the speaker-source

Speaker-source

Iconic gesture: depicting a round object with both hands.

(a) **(b)**



una caldera, es decir, un recipiente donde se calentaba agua
(a) **(b)**

a boiler, that is, a recipient where water has heated

Interpreter 2

Iconic gesture: depicting a round object with both hands.

(c) **(d)**

une chaudière à cuire
(c) **(d)**

a heater to cook

Figure 3. Interpreter 2 uses the same gesture type as the speaker-source

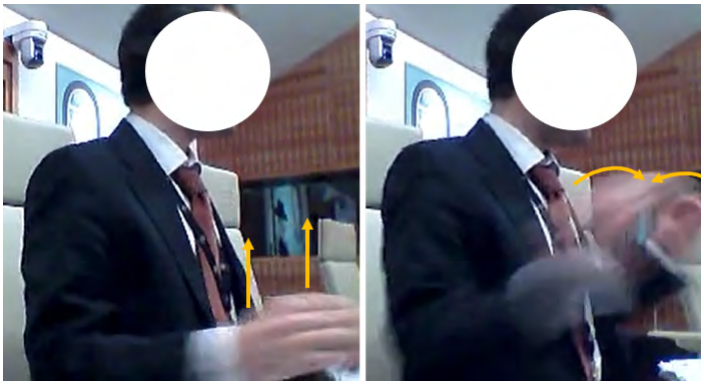
Incipient gestural alignment (+) — Interpreter gestures, but uses a gesture of a different type than that of the speaker

Speaker-source

Metaphoric gesture: first lifting, then moving both hands to the center, as if putting two things together.

The gesture is performed when mentioning the possibility of two persons taking care of the same activity.

(a) **(b)**



que uno baje la basura y otro se ocupe del lavaplatos
porque si los dos se ocupan del lavaplatos
(a) **(b)**
hay eh va a haber un problema

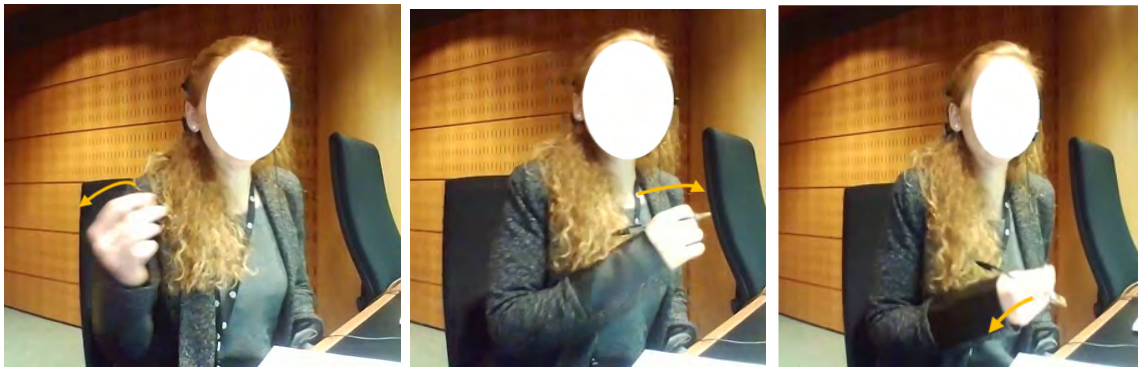
one should take out the trash, and the other should take care of the dishwasher, because if both take care of the dishwasher, there will be a problem

Interpreter 2

Deictic gesture (pointing forward to new referent, 'both') (caption e).

Previously, she used the same hand to point alternatively rightwards (caption c) and leftwards (d), when referring to the distribution of tasks between two subjects.

(c) **(d)** **(e)**



qu'un sorte les poubelles et que l'autre vide la vaisselle
(c) **(d)**
parce que si les deux s'occupent du lave vaisselle
(e)

one should take out the trash, and the other should take care of the dishwasher, because if both take care of the dishwasher

Figure 4. Interpreter 2 uses a gesture of a different type

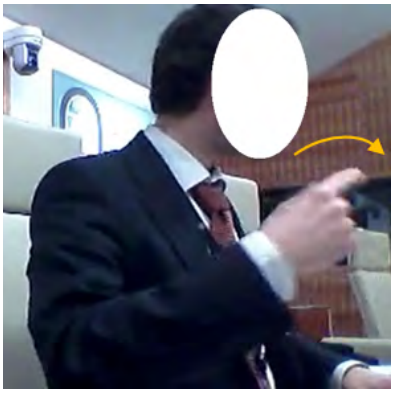
<table>
<tr><td colspan="2">No gestural alignment (×) — Speaker-source gestures but interpreter does not gesture</td></tr>
<tr><td>Speaker-source
Deictic gesture (pointing forward, new topic) with another person.
Previously, the same hand had pointed leftward, when saying treinta y seis años antes, ‘thirty-six years earlier’.</td><td>Interpreter 1
No gesture with another person.
Previously, her right hand points to her right when saying thirty-six years previous. After that, she holds the pen with both hands in a resting position and keeps them that way for the rest of the sequence.</td></tr>
<tr><td>(a)
  </td><td>(b)
  </td></tr>
<tr><td>treinta y seis años antes otra persona ya había patentado una máquina
(a)
thirty-six years earlier, another person had already patented a machine</td><td>thirty-six years previous to this another person had tried to patent a similar machine
(b)</td></tr>
</table>

Figure 5. Speaker-source gestures but Interpreter 1 does not gesture

3. Results

A descriptive quantification of the interpreters’ performance and a basic quantitative comparison of their behavior and that of the speaker-source are offered below in Tables 4 and 5.

	Interpreter 1		**Interpreter 2**	
	Hits	Rate	**Hits**	Rate
Speech interpreted—same type of gesture	50	42.4%	53	44.9%
Speech interpreted—different type of gesture	22	18.6%	27	22.9%
Speech interpreted—no gesture	28	23.7%	28	23.7%
Speech not interpreted	18	15.3%	10	8.5%
Speaker-source baseline →	**118**	**100%**	**118**	**100%**

Table 4. Overview of interpreters’ gestural performance

With regard to the interpreters’ overall performance (Table 4), two main tendencies should be noted. First, there are quite a few cases where the original speech of the speaker-source was not interpreted (15.3% of the cases for Interpreter 1; 8.5% for Interpreter 2). This is not surprising, as both are novice interpreters who sometimes experience disfluencies and miss or skip certain chunks of the speaker’s discourse. Interpreter 1 had more disfluencies or missed more speech sequences than Interpreter 2 (18 to 10), which might reveal an overall lower degree of interpreting competence.

At any rate, as second major tendency, there are no substantial differences in the degrees of gestural alignment exhibited by both interpreters in the sequences that were interpreted.

Furthermore, around 40% of the gestures by the speaker-source were mimicked by the interpreters through a gesture of the same kind.

	Baseline: gestures by speaker-source	Nr of hits where speech is interpreted along with *same type* of gesture				Nr of hits where speech is interpreted along with a gesture *of any type*			
		Interpreter 1		Interpreter 2		Interpreter 1		Interpreter 2	
Gesture type	**Hits**	**Hits**	**Rate***	**Hits**	**Rate***	**Hits**	**Rate***	**Hits**	**Rate***
Deictic (discourse structure)	40	13	32.5%	20	50%	20	50%	25	62.5%
Beat	24	10	41.6%	11	45.8%	15	62.5%	14	58.3%
Iconic	24	17	70.8%	14	58.3%	17	70.8%	22	91.6%
Metaphoric	24	9	37.5%	8	33.3%	16	66.6%	17	70.8%
Adaptor	5	0	0%	0	0%	3	60%	2	40%
Head shake (negation)	1	1	100%	0	0%	1	100%	0	0%
	118	**50**		**53**		**72**		**80**	

*Percentage of the speaker-source's gestures that prompt a gesture by the interpreter.

Table 5. Speech & gesture hits by the speaker-source that are interpreted along with a gesture

Table 5 reflects the gestural response of the interpreters according to gesture types by the speaker-source. As shown there, a total of 72 gestural responses by Interpreter 1 and 80 for Interpreter 2 were identified and classified. Looking at the hits by the speaker-source that receive a maximally aligned gestural response (gesture of the *same type*), iconic gestures clearly stand out as the type that more often elicit a mimicking response by the novice interpreters. It should also be noted that, although triggering a gestural response *of any type* by both interpreters in at least half of the cases (even in 60-70% in Interpreter 1), beats and metaphoric gestures get a maximally aligned response (*same gesture type*) in a smaller proportion, this reduction being much clearer for metaphoric gestures.

With regard to the gesture types that more often trigger *any kind* of gesture by the interpreters, iconic, metaphoric and beat gestures again stand out in frequency in the response by both interpreters. Moreover, iconic gestures trigger a very high gestural response by Interpreter 2 —in 91.6% of the cases, she gestures as a response to iconic gestures.

When comparing the performance of both novice interpreters, convergences and divergences arise at different levels, with a pattern that is not clearly identifiable. The greatest similarities are primarily observed in the frequency with which both interpreters respond to metaphoric and beat gestures, whether with a gesture of the same type or any other kind of gesture. The current sample size limits, though, the possibility of conducting statistical analyses of significant differences between the two interpreters. A substantial expansion of the analyzed data will allow for such a study in the future.

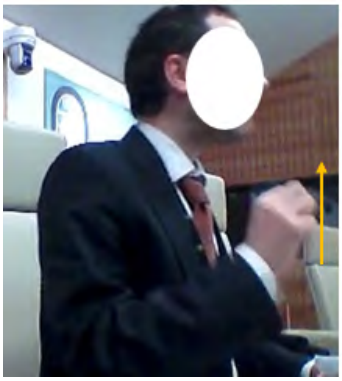
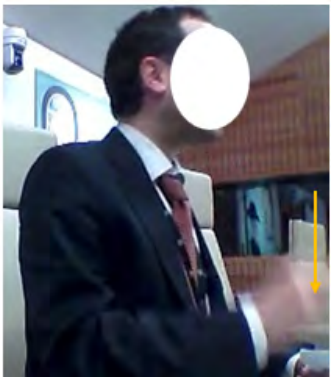
4. Discussion and conclusions

The results in section 3 allow to assess the research questions and hypotheses that were formulated above in section 2.2.

> **Research question 1.** Do the different gesture types by the speaker-source prompt diverse degrees of alignment by the interpreters in our data?
>
> > ***Hypothesis 1.*** *Gestures that do not relate to the speech content (beats, self-adaptors) prompt lower degrees of gestural alignment by the interpreters.*

The first hypothesis is not confirmed in our study, as it has been shown that sequences with beats are interpreted with gestures of any type in a good number of instances (62.5%, in Interpreter 1; 58.3%, in Interpreter 2), with still a notable proportion of cases where they are replicated by beat gestures (41.6%, in Interpreter 1; 45.8%, in Interpreter 2). Self-adaptors rarely appear in the sample analyzed in this paper and, consequently, their relationship to gestural alignment cannot be properly assessed.

However, the results obtained for beat gestures suggest that, in simultaneous interpreting, gestural alignment may not necessarily be driven by the distinction between semantic and non-semantic gestures (Ekman & Friesen, 1972; Kendon, 2004), that is, between gestures related or not related to the speech content with what is actually conveyed by words. In this vein, it could be claimed that simultaneous interpreting is of course guided by the semantic common ground that interpreters 'negotiate' with the speaker-source and the audience (*grounding views of alignment*), but *also* by other features of the linguistic behavior of the speaker-source, such as speech rhythm, prosody, and the beat movements that go along with them. Such a claim might —at least partly— support the *priming approaches to alignment* in a complementary and more comprehensive understanding of the coupling processes that regulate simultaneous interpreting. An example of an especially prominent alignment across all these aspects (rhythm, prosody, gesture) is provided in Figure 6. In this case, the speaker-source performs three beat gestures with his right hand when citing the title of a magazine section ('How to save your relationship'). These beats serve the function of parsing and stressing a segment of reproduced discourse (the section title). The title is cited verbatim, and so the hand also takes the form of a 'precision grip' gesture, as described by Kendon (2004, pp. 225-228). In Figure 6, the execution of the first of these beats by the speaker-source is visually depicted, with very broad and visible preparation and stroke phases. The interpreters' responses exhibit alignment on multiple levels: not only verbally, with a similar citation of the magazine section title, but also gesturally and in terms of rhythm and prosody, as the beat gestures they also perform with their right hands are synchronized with the same parts of the speech, emphasizing the quoted nature of the segment they accompany. Moreover, although both interpreters are holding a pen, the shape of their hands in some of their beats is compatible with a 'precision grip' gesture that serves the function of rhythmically parsing a segment of reproduced literal discourse.

Speaker-source	
en una sección de esta revista llamada **así salva usted** **beat 1 beat 2** su **relación** **beat 3** *in a section of this magazine called "how to save your relationship"*	**Beat 1 (preparation)**  **Beat 1 (stroke)** 

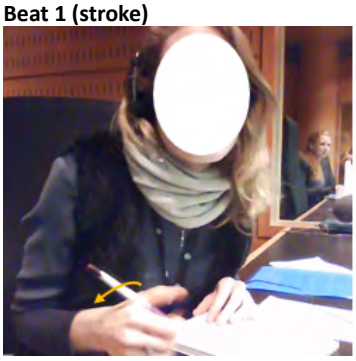
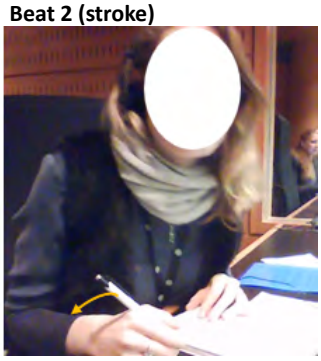
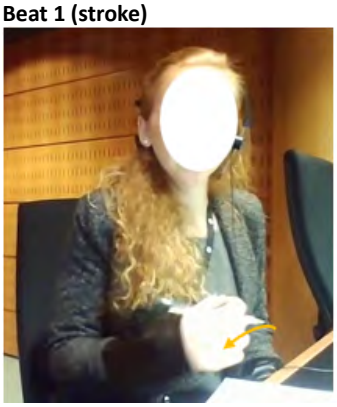
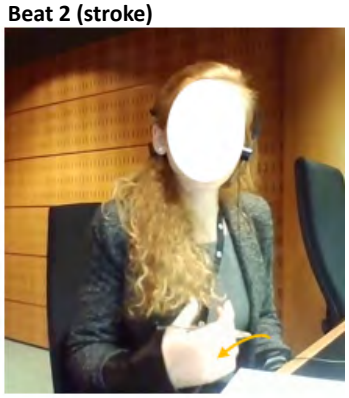
Interpreter 1	
in a section of the magazine how to **save** your re**lation**ship **beat 1** **beat 2**	**Beat 1 (stroke)**  **Beat 2 (stroke)** 
Interpreter 2	
dans la rubrique **comment sauver** sa **relation** **beat 1** **beat 2** **beat 3** *in the section "how to save your relationship"*	**Beat 1 (stroke)**  **Beat 2 (stroke)** 

Figure 6. Priming: alignment in gesture, rhythm and prosody

Hypothesis 2. *Within the realm of gestures relating to the representational (referential) or pragmatic (metadiscursive) meaning of speech, iconic gestures and discourse structuring gestures prompt higher degrees of gestural alignment.*

The hypothesis is confirmed only for iconic gestures, which clearly are the gesture type that is connected with a higher degree of gestural alignment on a more frequent basis and across both novice interpreters. This result confirms previous evidence in the same direction (Chwalczuk, 2021; Olza, 2024). As for deictic gestures with discourse structuring functions, results show a lower but notable triggering capacity for them, especially in Interpreter 2, who gestures in response to 62.5% of the cases, and replicates the same kind of gesture in 50% of the instances. Beat gestures seem to behave in a similar way to discourse structuring gestures, though. Therefore, our study is not conclusive on the operativity of the representational/ non-representational distinction (referential vs pragmatic gestures), nor the (non-)semantic one (beats and self-adaptors vs the rest of gesture types), to tackle gestural alignment in simultaneous interpreting.

Hypothesis 3. *Compared to other types of representational gestures, metaphoric gestures prompt lower degrees of gestural alignment.*

The hypothesis is not confirmed when the mere presence/absence of gesture by the interpreter is tracked, as metaphoric gestures follow iconic gestures in prompting a gestural response by the interpreters (66.6%, Interpreter 1; 70.8%, Interpreter 2). That being said, metaphoric gestures do exhibit more difficulties to elicit a maximally aligned response through another metaphoric gesture. The percentages reduce to 37.5% (Interpreter 1) and 33.3% (Interpreter 2) when looking at responses with the same type of gesture. As Leonteva et al. (2023, pp. 830-831)

claim, the production of metaphoric gestures involves the depiction and projection of actions in a physical domain (e.g., holding, molding, tracing, etc.) into an abstract domain, making it cognitively more demanding than performing non-metaphoric gestures, where only one representational domain is addressed. Regarding our results, and in line with these authors, it can be posited that, although metaphoric gestures may function as effective gesture primers, the demands of interpreting tasks make it difficult for interpreters to maintain the same level of metaphoricity in their gestures, leading them to use non-metaphoric gestures (e.g., iconic) in response to the speaker-source.

> **Research question 2.** Does gestural alignment rely on individuals? Or, by the contrary, does it work similarly in both interpreters who were observed?
>
> > ***Hypothesis 4.*** *The degree of gestural alignment exhibited by both interpreters is different due to personal styles and/or differences in fluency and performance quality.*

The results are inconclusive. On the one hand, differences for both interpreters were attested in fluency and competence (Table 4). Furthermore, the breakdown of their performance according to gesture types (Table 5) reveals some divergences in their gestural response the speaker source: in general terms, Interpreter 2 seems to respond more often to all kinds of gestures. In contrast, the two interpreters coincide in at least three main trends: they appear to be more sensitive to iconic gestures by the speaker-source; they also respond in notable ways to beats and metaphoric gestures; and they prefer other gesture types when aligning with metaphoric gestures. In our previous approach to another sample from the same dataset (Olza, 2024), clearer differences between the interpreters were observed. For instance, Interpreter 1 exhibited a much lower percentage of non-gestural hits –that is, of cases where the interpreted sequence was not accompanied by a gesture— compared to Interpreter 2 (7.3% for Interpreter 1; 25% for Interpreter 2). Although in the present study the performance of both interpreters was found to be more similar, a future analysis of the entire dataset will allow for proper statistical tests to better delineate the differences in their performance.

All in all, the most relevant findings of this study can be summarized in two directions. In the first place, the two interpreters under observation maximally aligned with the speaker-source at the gestural level, using the same type of gesture, in around 40% of the cases. Also, they gesturally responded to the speaker-source—irrespective gesture types—in more than 60% of the instances. To sum up: in our data, gestural alignment is more a norm than an exception. In the second place, iconic gestures were the gesture type that more often and better prompted gestural alignment by the interpreters. Beats and metaphoric gestures also elicited notable degrees of alignment.

A study like the one offered here shows that gestural alignment in simultaneous interpreting is still to be explored and understood in several uncharted territories. The results explained above nevertheless stress an uncontested claim in the field, which is that empirical evidence on interpreting tasks does not fit the ‘conduit model’, that is, it shows that interpreters do not merely transfer meanings from one language to another, mechanically decoding what the speaker says and then recoding it in exactly the same way in the target language, as described in the ‘conduit metaphor’ for language (Reddy, 1979, pp. 286-292), which was critically reviewed by Reddy himself in his seminal work (1979, pp. 297-310). Instead, their performance is better explained through a model that integrates the complex set of cognitive, linguistic, and behavioral conditions that influence the interpreters’ activity, which is more of a cooperative task than merely an ‘imitative’ one (Janzen et al., this special issue).

To further advance in the understanding of this complex set of factors and effects, a study like

this one will need to develop in several directions. For instance, the verbal response of the interpreters is still to be systematically studied, to analyze how the linguistic choices they make affect their own gestural behavior. In addition, a more thorough formal analysis of the gestures by both the speaker-source and the interpreters would allow to refine the conclusions offered here, as even the cases of what we have here considered as 'maximal gestural alignment' (same gesture type by the interpreters) exhibit interesting differences in the material articulation of the body movement, with different imagery and interpersonal features involved in them. Finally, other limitations of this study could be overcome with significantly broader data, as well as data even more closely aligned with the reality of professional interpreters. This would include gathering data from more experienced interpreters engaged in tasks that more accurately reflect their actual practice. As has been noted, the data for this study comes from training exercises with novice interpreters in a real courtroom setting, but in tasks different from strict legal interpretation. Therefore, it remains necessary to gather and analyze audiovisual data from experienced interpreters who either align or do not align gesturally with the speaker in real courtroom sessions.

In spite of its limitations, this study decidedly supports the call for a more multimodally oriented research on, and training of, simultaneous interpreters (Salaets & Brône, 2020). Videos and multimodal data are indeed the key to fully integrate the gestural dimension into the analysis of simultaneous interpreting. And this will, in turn, lead to a better awareness of the importance of multimodality in the interpreters' own professional performance.

5. Acknowledgements

This research was funded by the Spanish Ministry of Science and Innovation and FEDER/EU, grant numbers PGC2018-095703-B-I00 (MultiNeg) and PID2022-143052NB-I00 (MultiDeMe), and the Institute for Culture and Society, grant number ICS2020/09 (InMedio). Also, it was developed within the frame of the CoCoMInt Research Network of the Spanish Ministry of Science and Innovation (ref. RED2022-134123-T).

The author wants to thank the interpreters who generously agreed to be recorded for this study, and those who facilitated the fieldwork and recordings.

6. References

Bergmann, K. & Kopp, S. (2012). Gestural alignment in natural dialogue. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, *34*, 1326-1331. https://escholarship.org/uc/item/73z0q063 (accessed on 15 January 2024).

Bressem, J. & Müller, C. (2014). The family of away gestures: Negation, refusal, and negative assessment. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & S. Tessendorf (Eds.), *Body—language—communication. An international handbook on multimodality in human interaction* (pp. 1592-1604). Mouton de Gruyter. https://doi.org/10.1515/9783110302028.1592

Chwalczuk, M. (2021). *La gestualité co-verbale en interprétation dans les services publics: Analyse contextualisée d'un corpus multimodal*. Doctoral dissertation, Université Paris Cité, France. https://theses.hal.science/tel-03509754/document.

Cienki, A. (2022). The study of gesture in cognitive linguistics: How it could inform and inspire other research in cognitive science. *WIREs Cognitive Science*, *13*(6), 1-17. https://doi.org/10.1002/wcs.1623

Cienki, A. (2025). Functions of gestures during disfluencies in simultaneous interpreting of speech, *Parallèles*, *37*(1), 29-46. https://doi.org/10.17462/PARA.2025.01.03

Doyle, G. & Franck, M. C. (2016). Investigating the sources of linguistic alignment in conversation. In K. Erk & N. A. Smith (Eds.), *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (pp. 526-536). Association for Computational Linguistics. https://doi.org/10.18653/v1/P16-1050

Eilan, N., Hoerl, C., McCormack, T. & Roessler, J. (Eds.) (2005). *Joint attention: Communication and other minds*. Oxford University Press.

Ekman, P. & Friesen, W. V. (1972). Hand movements. *The Journal of Communication, 22*(4), 353-374. https://doi.org/10.1111/j.1460-2466.1972.tb00163.x

Fernández Santana, A. & Martín de León, C. (2022). Con la voz y las manos: Gestos icónicos en interpretación simultánea. *HERMĒNEUS,* 23, 225-271. https://doi.org/10.24197/her.23.2021.225-271

Fusaroli, R. & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science, 40*(1), 145-171. https://doi.org/10.1111/cogs.12251

Giles, H., Coupland, N. & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1-68). Cambridge University Press.

Giles, H. & Ogay, T. (2007). Communication Accommodation Theory. In B. B. Whaley & W. Samter (Eds.), *Explaining communication: Contemporary theories and exemplars* (pp. 293-310). Routledge.

Goodwin, C. (2018). *Co-operative action*. Cambridge University Press.

Healey, P. G. T. (2004). Dialogue in the degenerate case? *Behavioral and Brain Sciences, 27*(2), 201. https://doi.org/10.1017/S0140525X04330050.

Iriskhanova, O., Cienki, A. J., Tomskaya, M. V. & Nikolayeva, A. I. (2023). Silent, but salient: Gestures in simultaneous interpreting. *Research Result: Theoretical and Applied Linguistics, 9*(1), 99-114. https://doi.org/10.18413/2313-8912-2023-9-1-0-7

Janzen, T., Leeson, L., & Shaffer, B. Simultaneous interpreters' gestures as a window on conceptual alignment. *Parallèles*, *37*(1), 85-104. https://doi.org/10.17462/PARA.2025.01.06

Kendon, A. (2004). *Gesture. Visible action as utterance*. Cambridge University Press.

Kimbara, I. (2006). On gestural mimicry. *Gesture, 6*(1), 36-61. https://doi.org/10.1075/gest.6.1.03kim

Kopp, S. & Bergmann, K. (2013). Automatic and strategic alignment of co-verbal gestures in dialogue. In I. Wachsmuth, J. de Ruiter, P. Jaecks & S. Kopp (Eds.), *Alignment in communication: Towards a new theory of communication* (pp. 87-108). John Benjamins.

Krauss, R. M. & Pardo, J. S. (2004). Is alignment always the result of automatic priming? *Behavioral and Brain Sciences*, *27*(2), 203-204. https://doi.org/10.1017/S0140525X0436005X

Krystallidou, D. (2020). Going video: Understanding interpreter-mediated clinical communication through the video lens. In H. Salaets & G. Brône (Eds.), *Linking up with video: Perspectives on interpreting practice and research* (pp. 81-202). John Benjamins.

Leonteva, A. V., Cienki, A. & Agafonova, O. V. (2023). Metaphoric gestures in simultaneous interpreting. *Russian Journal of Linguistics*, *27*(4), 820-842. https://doi.org/10.22363/2687-0088-36189

Martín de León, C. & Fernández Santana, A. (2021). Embodied cognition in the booth: Referential and pragmatic gestures in simultaneous interpreting. *Cognitive Linguistic Studies*, *8*(2), 277-306. https://doi.org/10.1075/cogls.00079.mar

McNeill, D. (2005). *Gesture and thought*. The University of Chicago Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. The University of Chicago Press.

Olza, I. (2024). Gestural alignment in spoken simultaneous interpreting: A mixed-methods approach. *Languages, 9*(4), 151. https://doi.org/10.3390/languages9040151

Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169-190. https://doi.org/10.1017/S0140525X04000056

Rasenberg, M., Özyürek, A. & Dingemanse, M. (2020). Alignment in multimodal interaction: An integrative framework. *Cognitive Science*, *44*(11), e12911. https://doi.org/10.1111/cogs.12911

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284-324). Cambridge University Press.

Riordan, M. A., Kreuz, R. J. & Olney, A. M. (2014). Alignment is a function of conversational dynamics. *Journal of Language and Social Psychology*, *33*(5), 465-481. https://doi.org/10.1177/0261927X135123

Salaets, H. & Brône, G. (Eds.) (2020). *Linking up with video: Perspectives on interpreting practice and research*. John Benjamins.

Stachowiak-Szymczak, K. (2019). *Eye movements and gestures in simultaneous and consecutive interpreting*. Springer. https://doi.org/10.1007/978-3-030-19443-7

Stivers, T. (2008). Stance, alignment and affiliation during storytelling: When nodding is a token of preliminary affiliation. *Research on Language in Social Interaction*, *41*(1), 29-55. https://doi.org/10.1080/08351810701691123.

Stivers, T., Mondada, L. & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada & J. Steensing (Eds.), *The morality of knowledge in conversation* (pp. 3-24). Cambridge University Press.

Zagar Galvão, E. (2013). Hand gestures and speech production in the booth: Do simultaneous interpreters imitate the speaker? In C. Carapinha & I. Santos (Eds.), *Estudos de lingüística* (pp. 115-129). Coimbra University Press.

Zagar Galvão, E. (2015). *Gesture in simultaneous interpreting from English into European Portuguese: An exploratory study*. Unpublished Doctoral dissertation, University of Porto, Portugal.

Zagar Galvão, E. (2020). Gesture functions and gestural style in simultaneous interpreting. In H. Salaets & G. Brône (Eds.), *Linking Up with video: Perspectives on interpreting practice and research* (pp. 151-180). John Benjamins. https://doi.org/10.1075/btl.149.07gal

Inés Olza
University of Navarra
Institute for Culture and Society
Edificio Ismael Sánchez Bella
31009 Pamplona
Spain

iolzamor@unav.es

Biography: Inés Olza is Senior Researcher in Language and Cognition at the Institute for Culture and Society (ICS) of the University of Navarra, Spain. Within ICS, she is Principal Investigator of the Bonds, Creativity and Culture Research Group, and director of the Multimodal Pragmatics Lab (MuPra Lab). Her research focuses on figurative language, gesture, phraseology and interactional dynamics from the perspective of Pragmatics, Cognitive Linguistics and Multimodality.

Parallèles 37(1) April 2025

Simultaneous interpreters' gestures as a window on conceptual alignment

Terry Janzen[1], Lorraine Leeson[2], Barbara Shaffer[3]

[1]University of Manitoba, [2]Trinity College Dublin, [3]University of New Mexico

Abstract

In this qualitative study we examine relationships between gestural alignment and conceptual alignment between speakers' source texts and simultaneous interpreters' target texts. We find that interpreters' gestures provide a window into their conceptualizations of source text elements, but also that gestural-conceptualization relationships are complex. We report on spoken-to-spoken language data, taken from a larger study, where interpreters interpreted from English to French, Spanish, Navajo, and Ukrainian. Each interpreter was video-recorded interpreting two English texts into their target language, followed by a video-recorded Stimulated Recall where they discussed whether their visualizations of the source text aided in how they understood the text. We find evidence of multi-level cognitive blends, where the interpreter's own subjective experiences blend with their assessment of the speaker's viewpoint, rather than the interpreter fully assuming the speaker's viewpoint. The data reveal instances of gestural alignment and corresponding conceptual alignment, gestural and conceptual non-alignment, and less-clear cases that suggest a complex relationship between gesturing and conceptualization. As a result, we propose a typology of gestural and conceptualization alignment/non-alignment in the interpreters' target texts.

Keywords

Conceptual alignment, gestural alignment, visualization, subjectivity, simultaneous interpreting

 DOI:10.17462/para.2025.01.06

1. Introduction

Simultaneous interpreters are primarily concerned with speakers' meaning, and conveying some sense of this meaning between the speaker and recipient, as is evident throughout the literature on interpretation and translation from various perspectives (see among many others Nida, 1964; Gerver, 1976; Seleskovitch, 1978; Gile, 1995; Hatim & Mason, 1990; Wadensjö, 1998; and Setton, 1999). From a cognitivist perspective, meaning resides in the mind, and because we have no direct access to another's mind, meaning must be constructed by the addressee based on available clues via what the speaker says, along with prosodic features of their discourse, their gestures, etc. (Croft, 2000; Linell, 2009; Reddy, 1993). Given that what the speaker says is comprised of words and constructions that are but representations of the meaning in their mind, the addressee is largely dependent on their own process of assembling a meaning, filtered through their own experiences both of how things "are" (i.e., encyclopedic knowledge) and of language and what words and constructions typically mean in contextualized settings (Croft, 2000).

This project sets out to explore one aspect of this enterprise, that is, the role that gesture plays in how the simultaneous interpreter interacts with the source text (Zagar Galvão, 2013), in two ways: 1) how the interpreter engages with a speaker's gestures as part of their delivery of a text, and 2) how the interpreter's own gestures reflect elements of their conceptualization of the text meaning, and how this may influence the resulting target text – the text that the interpreter delivers to the target audience. Taylor (2017) suggests that in the process of translation, the translator pictures a situation as the original writer depicts it, but then considers how this conceptualization can be best represented in the target language, given that the language resources for doing so may not allow for a depiction in the same way in the target as is the case for the source text. We were interested in whether the interpreters' conceptualization of source text elements aligned with source speakers' own conceptualizations of these elements, and hypothesize that the interpreters' own gestures offer a window on this alignment. To discover this, we analyzed the interpreters' enactment gestures (Ferrara & Johnston, 2014; Saunders & Parisot, 2023), often referred to as depicting (Liddell & Metzger, 1998), and their deictic gestures (see for example Kita, 2003).

1.1. The interpreter's inherent subjectivity

Despite the ideology of the interpreter striving toward unbiased objectivity, a cognitivist view asserts that speakers' participation in discourse is inherently subjective at every level, and interpreters are not exempt from their own inherently subjective approach to the discourse. Afterall, at even the basic level of linguistic expression, it is the interpreter's "words, her grammar, her intonation and prosody, and her set of experiential frames that she has been building, all of which conflate in her use of language that becomes the target text" (Janzen & Shaffer, 2013, p. 79). It is well understood that the interpreter filters incoming texts through their own experiential, subjective understanding of the world (Janzen & Shaffer, 2008; see also Boogaart & Reuneker, 2017; Linell, 2009), and this includes cognitive resources[1] and subjective conceptualizations of the world in constructing meaning. Critical is that such conceptualizations are subjectively viewpointed (Sweetser, 2023) and dynamic (Langacker, 2008).[2] This leads to

[1] Janzen (2005) frames linguistic form, text building strategies, and even text meaning (sense) as among the interpreter's resources in constructing a target text, highlighting the subjective nature of target text construction.

[2] Dynamicity, in Langacker's (2008) terms, suggests that conceptualizations are always subject to new information, and therefore subject to change; therefore "conceptualizations" as a term is preferred to "concepts", which implies something static, not reflective of actual discoursal cognition (e.g., Linell, 2009).

the potential of the interpreter's target text aligning conceptually with the source speaker or not. If the interpreter is inextricably bound by their own conceptualizations of the world as they see it and their subjective conceptualizations of the source text, is there any possibility of an objective representation of the source text at all? Nonetheless, interpreters are often under the impression that they are to represent and portray the viewpoint of the source speaker, and are expected to do so in an objective way (see, e.g., Tipton, 2008; Wadensjö, 1998; Wilcox & Shaffer, 2005 on interpreter neutrality; and Venuti, 1995 on translator invisibility). In a preliminary analysis of the data in this project, Leeson et al. (2017a, 2017b) and Janzen et al. (2022) find that interpreters form a blended viewpoint that is partly their own viewpoint on speaker meaning, and partly their subjective *belief* of what the speaker's viewpoint is. Critically, the understanding of subjective belief is a departure from the idea that an interpreter's rendition is an actual representation of the speaker's viewpoint.

Underlying the interpreter's blended viewpoint is that some sort of mental simulation is taking place, where the interpreter experiences mental re-enactments of sensory-motor states (Barsalou, 2003; Cienki, 2013) as described by the source speaker. In discourse studies, Bergen (2005, p. 262) describes this as "simulation semantics" wherein understanding what a speaker is saying entails "performing mental perceptual and motor simulations" of the text.

1.2. Situating gestures in simultaneous interpreting

Mental simulation as described above can involve body actions that are gestural in that they are communicative, intentionally or not, given that gestures may be intended as interactive so as to communicate something to an addressee or are reflexive in the sense that they appear to assist the speaker in conceptual processing or in word recall (Frick-Horbury, 2002; Kita et al. 2017; see also the review in Cooperrider & Goldin-Meadow, 2017). This would be the case for referential gestures that concern content within the discourse, and gestures that reflect an attitude toward the content or indicate to the recipient how they might understand content framing (Cienki, 2024). Pragmatic gestures in particular can reflect stance taking (Leonteva et al., 2023). Wu and Coulson (2007, p. 244) suggest that "iconic gestures activate image-specific information about the concepts which they denote", which could be the case both for the speaker/gesturer and the addressee. Sweetser (2023) makes the critical point that the meaning of a gesture must be considered within the context of a *viewpointed* gesture space, with the gesture not considered as an isolated action of a body part. The body, therefore, is fundamental to embodied cognition and the situated meaning of gestures. This suggests that gestures are fundamental to mental simulation and, we argue, following Wilcox and Shaffer (2005) and Janzen and Shaffer (2013), interpreters cannot avoid this deeply embodied aspect of interactive discourse.

2. Description of the visualization project

The current analysis is part of a larger study on simultaneous interpreting that investigates the extent that interpreters working into either spoken or signed languages visualize aspects of the source text, and how such visualizations inform the interpreter's construction of meaning and their decision-making processes in building their target text. In addition, given that an important characteristic of discourse is that it is multimodal (among many others, Enfield, 2009; Hagoort & Özyürek, 2024; Sweetser, 2023), we wanted to explore whether the interpreters were cognizant of the source speakers' gestures, and whether they considered these gestures as contributing to speaker meaning, therefore incorporating this information into their target text construction. We were also interested in the interpreters' own gestures as they interpreted, considering that they may reveal elements of conceptualizations not

apparent in the spoken or signed components[3] of the interpreters' utterances (Janzen et al., 2023). It is these gestural elements and how they align with conceptualization that concern this present analysis, focusing on the spoken-to-spoken language interpreters in the study. Olza (2024, this special issue) also examines gestural alignment in terms of gesture type between the simultaneous interpreter and source speaker, but her analysis does not focus on gestural-conceptual alignment.

Our participants are fourteen professional interpreters with a minimum of five years of experience in simultaneous interpreting, working between English and French, Spanish, Ukrainian, Navajo, American Sign Language (ASL), and Irish Sign Language (ISL). Each participant was asked to interpret two spoken English texts: an interview with the Canadian astronaut Chris Hadfield (approximately 15 minutes)[4] and a segment of a live performance of the Irish comedian Dara O'Briain (3.45 minutes)[5]. Data collection took place in Canada, the US, and Ireland, and followed university ethics guidelines in each country. Consent was given by all study participants for their images and video clips to be used in public presentations on the study, and in published reports. The participants were first given the opportunity to watch both source text videos, and then were video-recorded interpreting the two texts. Immediately following this, we recorded a Stimulated Recall (SR) (Bloom, 1954; Russell & Winston, 2000) where we reviewed, post-task, the interpretation while viewing both the source text video and the interpretation, as shown in Figure 1. SRs represent a methodology in interpreting research that affords researchers insights into interpreters' cognitive processing during their interpretations. Think Aloud Protocols (TAPS) on the other hand involve reporting concurrent with translation or other activities. See, for example, the studies in Tirkkonen-Condit and Jääskeläinen (2000), and Russell and Winston (2014) on SRs, also referred to as retrospective process tracing (Herring & Tiselius, 2020).

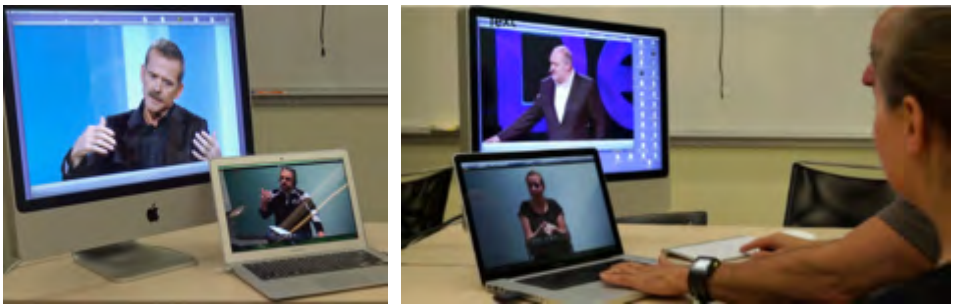


Figure 1. The SR setup: viewing the source text and simultaneous interpretation together

The participants were told only that we were collecting examples of simultaneous interpreting, thus were unaware that our focus was on their visualization strategies along with the source speakers' and participants' gestures during the simultaneous interpreting task.

[3] While data on the signed language interpreters in the study are not included here, the question of distinguishing what is "gestural" and what is "linguistic" has been a topic of debate for both signed and spoken language. Some theoretical perspectives consider the production of speech sounds to be gestural, for example the work on articulatory phonology (Browman & Goldstein, 1992; Neisser, 1976). Nonetheless, details on this are beyond the scope of this paper; here we consider gestures to be body actions other than the articulation of speech itself.

[4] From an episode of *The Hour* (CBC, Canada), https://www.cbc.ca/strombo/videos/chris.hadfield-full-interview-strombo

[5] https://www.youtube.com/watch?v=BVxOb8-d7Ic

3. The alignment of conceptualization and gesture

As described above, we were interested in learning more about conceptual alignment between the source speaker and the interpreter, considering that such alignment would undoubtedly be the interpreter's goal, that is, constructing a meaning that would match the intended meaning of the speaker. Detailed analyses of the verbal interpretations regarding equivalence are left for future discussion; here our focus is on what the interpreters' gestures reveal about their conceptualizations of the ideas expressed in the source texts. To this end, we look both to the gestures accompanying the actual simultaneous interpretations and what the interpreters said in the SRs about how they conceptualized the texts, and occasionally, their gestures as they talked in the SRs about their conceptualizations, which at times were equally revealing, even if after the fact.

This study is a qualitative analysis. We have not yet, at this point, examined every relevant gesture in the videos, but report on specific examples that reveal aspects of conceptualization beyond the words the interpreters use either in their simultaneous interpretations or their SRs. In undertaking this study, we examined the video data with respect to both conceptual and gestural alignment along the following possible combinations:

i) conceptual alignment + gestural alignment
ii) conceptual non-alignment + gestural non-alignment
iii) conceptual alignment + gestural non-alignment
iv) conceptual non-alignment + gestural alignment

In what follows, we see clear examples of the first two possibilities, but found minimal evidence, at least in spoken language target texts, of the third and fourth possibilities (although we did find one example that seemed to qualify as exemplifying them in the signed language target text data; see section 3.3 below).

3.1. Conceptual alignment and gestural alignment

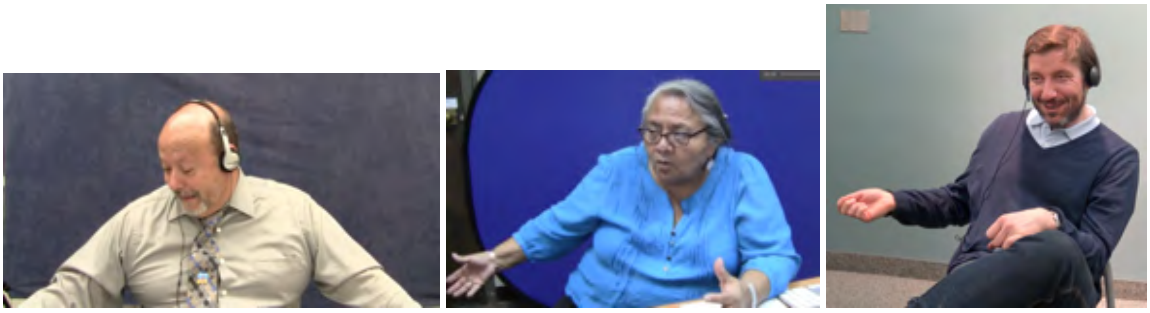
We found many instances where the interpreter's gesture(s) demonstrated alignment with the source speaker, indicating alignment in their conceptualization of the text. This is significant because in at least some cases, the speaker's gesture represented semantic or pragmatic information that was not part of the spoken text. In Example 1, the speaker, Dara O'Briain, asks what it would be like if some Renaissance figures suddenly appeared in our time, asking us to explain how common household things work, for example, electric appliances. O'Briain suggests that we don't really know how they work, or more exactly, how *electricity* works—you just plug them into the wall. The choice of direction of his pointing gesture to a low, distal rightward space, is significant. Janzen et al. (2023) show that in many discourse events, the gesture spaces that both speaker/gesturers and signed language users choose exemplify the conceptual metaphor CONCEPTUAL DISTANCE IS SPATIAL DISTANCE. Janzen et al. show that things that are known or knowable, including gestural references to past events and spaces, are referenced gesturally in a frontal proximal space that is within view of the speaker or signer, and maximally viewable to a face-to-face addressee. Gestural references to things not known or unknowable, including future events, things "out of mind", of an irrealis nature, even hypotheticals, occupy more distal gesture spaces, frequently out of range of a forward-facing visual viewpoint. In O'Briain's case, the unknowability of how electricity works is profiled by a wall positioned well away from an accessible frontal proximal space (Figure 2), and as shown in (2b), even staring at it will not help. Nothing about gesture space generally would prevent O'Briain from positioning the imaginary wall and socket directly in front of him. But while he does not mention why the wall is mentally positioned off to the side out of view, the effect it has for the viewer is clear conceptual alignment – they don't know how it works any

better than he does; even his gesture of looking at the space without saying anything (without accompanying speech) now garners audience laughter.

Example 1: The wall

(a) pointing (b) looking at

Figure 2. The source speaker (English) pointing to a low space to the far right in (a); subsequently turning to look at that space in (b)



(a) Spanish (b) Navajo (c) Ukrainian

Figure 3. Interpreters' gestures toward the rightward space: eye gaze in (a), and right-hand gestures in (b) and (c)

The gestures of some of our participants show close alignment with O'Briain's gestures as they interpreted these segments of the text (Figure 3). The Navajo interpreter demonstrated this alignment in the SR (3b), even though she had not made this gesture during the actual interpretation. None, however, reported consciously copying the speaker's gestures to this rightward distal space, and in the SR, some were quite surprised that they had even done so. Most interpreters simulated the event with the imaginary wall positioned in the same orientation, and continued to repeat this gestural orientation in the SR, highlighting its saliency for them. This suggests they understood the metaphoric sense of the speaker's gestures, which impacted their conceptualization of the overall sense conveyed by the speaker, that of an unknown mechanism (how electricity works), and as a result prompting their own gestures, as a reflection of the unknowable state. It should not go unnoticed the grins on the faces of the interpreters in Figure 3a and 3c, which suggest that rather than a fully enacted bewildered stance of the source speaker, they align with the audience reaction of hilarity, thus suggesting a body-partitioned (Dudis, 2004) portrayal of the event. On one level, the interpreter reflects and represents O'Briain's plea to ignorance of how things work, but this is overlayed with a representation of the staged performance and audience participation. Therefore careful examination of the interpreter's "performance" reveals a secondary, intersubjective reflection of audience response, simultaneously. This appears to be an example of what Dancygier (2012) refers to as stance-stacking.

In Example 2 Hadfield is talking about his experiences in space, including the perspective he gained of being able to see the planet Earth from space, so distant that he could "cover up the world with your thumb" as he says. Figure 4 shows his co-speech gesture in an enactment of raising one's thumb in the direction of the planet and looking directly at it, so as to imply that the planet is so far away, so small, as to be completely covered by the thumb. Hadfield's gesture here is a co-speech gesture, so that unlike Example 1 (and Example 3 below), the gesture meaning coincides with what is said, and is therefore not adding any meaning distinct from what is said.

As in the other examples, during the SRs the interpreters did not suggest that they were prompted to gesture in like manner because Hadfield had done so himself, and at best, some were only vaguely aware that he had made this gesture. Nonetheless, their gesturing at this point is striking. All 14 interpreters in the study gestured with their thumb in this way as they produced the target utterance. Thus, it is one of the clearest examples of gestural alignment. Conceptual alignment is evident as well, but it is interesting to consider, because none of the interpreters has had a similar experience to that of Hadfield's. However, by analogy, a common and relatable experience is one of covering the moon with your thumb, or covering any smaller, closer object with your thumb, and so being able to conceptualize such an enactment is not by any means a stretch for these interpreters, and the fact that all reproduced the gesture suggests that something particularly salient stood out for them. Of particular interest is the interpreter's physical stance in Figure 5(c). His eyes were closed during long segments of his interpretation, nonetheless, his body and head positioning clearly suggest a visual conceptualization of the act, illustrating our claim that the interpreters were not just seeing what Hadfield was doing and copying it.

Example 2: Covering the earth with your thumb

Figure 4. Hadfield gesturing while saying "being able to cover up the world with your thumb"

(a) Ukrainian (b) Spanish (c) French

Figure 5. Interpreters gesturing while giving the equivalent target text

Example 3 is more complex, both cognitively and gesturally. Here, the gestures represent abstract ideas with the gesture spaces signifying a discourse-level, cognitively-organizational act of comparison. Prior to this, Hadfield's interviewer focused on his work as an astronaut,

but at this point in the interview, he turns to ask about Hadfield's home and family life. Hadfield, who is righthanded, says, "these are my objectives at work (while gesturing to a contralateral leftward space), and these are my objectives at home (while gesturing to an ipsilateral rightward space)". The two sets of objectives, then, are referred to deictically with the demonstrative "these" but without specifying exactly what they are, differentiated by assigning them two distinct gestural spaces, as in Figure 6.[6]

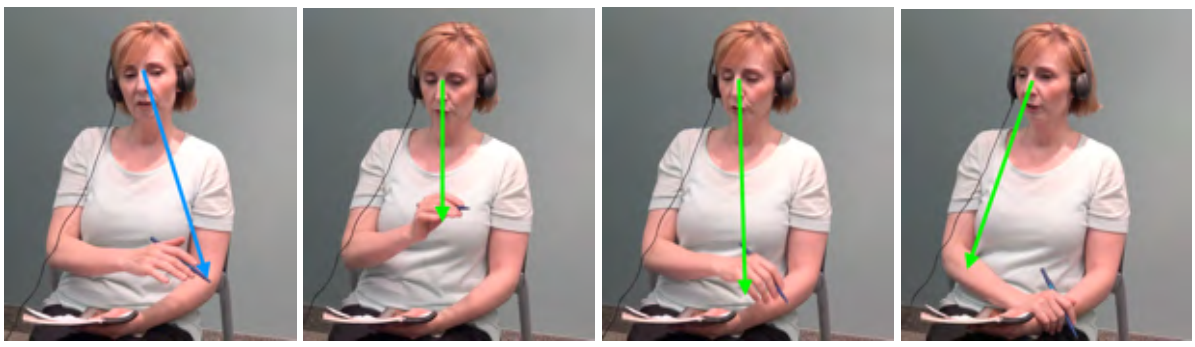
Many interpreters in the study gestured in a remarkably similar manner. The interpreter in Figure 7 shows the same gesture space differentiation, strengthened by her differentiated eye-gaze to the two locations. Figure 7(b-c) shows her arm lowering to the second space, continuing on to rest on her left wrist positioned (conveniently, perhaps) on that vertical plane (Figure 7d). In (7d) her eye-gaze has moved farther rightward, we assume because she is cognitively moving on to the next text item.

This gestural alignment is an index of conceptual alignment, but here, it is not alignment in terms of the content of the source text, but regarding the cognitive organization of ideas within a comparative frame. In the SRs, none of the interpreters suggested that they were aware of Hadfield's gestures to the two spaces and, when pointed out, none could articulate why they thought he might have done so. Even more interesting is that the participants who gestured in a similar way were surprised to see this in the recording. While we cannot rule out the possibility of unconscious mimicking (Chartrand & Bargh, 1999; Kimbara, 2006), it seems more likely that this metaphoric use of space to correspond with a conceptualized comparative frame is part of human cognitive architecture, as noted by others (e.g., Hinnell & Rice, 2016) and which has been considered as part of the "spatial grammar" of ASL (Winston, 1995).

Example 3: Comparing work and family at home

(a) work gesture space (b) family gesture space

Figure 6. Hadfield's differentiated gesture spaces for work (a), and home (b)



(a) (b) (c) (d)

Figure 7. A French interpreter's differentiation of gesture spaces for work in (a), and home in (b – d)

[6] Although Hadfield's hand dropped below the bottom of the video image, these stills show the beginning of his hand/arm movements, which clearly indicate the two distinct gesture spaces.

3.2. Conceptual non-alignment and gestural non-alignment

Just as gestural alignment can be a window into conceptual alignment as seen above, gestural *non*-alignment can signal conceptual non-alignment. Further, it appeared that at times this is unintentional, with the interpreter perhaps being unaware of it. On the other hand, there were instances where non-alignment was intentional. Examples of each are discussed below.

3.2.1. Unintentional conceptual non-alignment and gestural non-alignment

We found instances where the interpreter's gesture did not match the source speaker's, prompting our comparison of the source and target texts. Several things are apparent here. First, the interpreter may have misheard, or misunderstood what the speaker said, or could not come up with an equivalent in the target language. In Example 4, whereas the speaker says "they're little um, like, saloon doors", the Ukrainian interpretation is "Це як двері в салоні" ('it's like a cabin door'), and the gestures of closing the door(s) are different in both spatial orientation and pulling a single door shut instead of double saloon doors (Figure 8). We analyze instances like this as unintentional, as there is no reason to think that the interpreter is determining to say something other than what the speaker intends. In the SR this interpreter did not discuss his lexical choice of words ('cabin' instead of 'saloon'), but did describe his visualization of the sleep pod, noting that he was "in the zone", feeling comfortable in his interpretation. "Here, I am him," he commented. Alan Cienki (personal communication) suggests that двері is a "plurale tantum" in Ukrainian (as is 'scissors' in English), but it could translate into English as either 'door' or 'doors'. Throughout his interpretation of Hadfield, this interpreter frequently gestured with both hands, so his one-handed gesture here leads us to believe that it has been conceptualized as a single door.

Example 4: The saloon doors

Source text: "uh, pull the little doors closed, like 'thunk thunk', because they're little um, like, saloon doors, on your sleep pod, …"

Ukrainian target text: Можна закрити ... можна закрити двері. Це як двері в салоні. І це там де ти спиш. 'You can close ... you can close the door. It's like a cabin door. And this is where you sleep.' (transcription and English translation: Olena Gordiyenko)

(a) (b)

Figure 8. Hadfield's gesture of closing the "saloon" doors in (a); the interpreter closing a "cabin" door in (b)

3.2.2. Intentional conceptual non-alignment and gestural non-alignment

Conversely, we found several instances where neither a non-aligned gesture nor conceptual non-alignment were apparent in the interpretation but were revealed in the SR. The interpreters at times stated that they believed they were thinking quite differently from the source speaker, and it was here that the gestures were produced, and where the occurrence of intentional conceptual non-alignment came to light.

In Example 5, the Navajo interpreter gestured very little during the interpretation but much more so during the SR. In O'Briain's comedy sketch text he makes the claim to his imaginary Renaissance visitors that an advantage of modern technology is the ability to make two-sided toast without having to flip the bread. He gestures putting the bread in a top-slotted toaster and pushing down the lever (Figure 9a), dropping the bread down between the two heating elements. In the Navajo interpreter's SR, she explained that in her community and culture, her target audience would not understand if she tried to follow the source text because they would not have experience with such a toaster. Instead, she adopts a target audience perspective to make the target text maximally accessible, visualizing and describing making toast in a wire basket over an open fire (Figure 9b). In doing this, she sacrificed the joke but raised the level of meaning in the target text.[7]

Example 5: Two toasters

(a) (b)

Figure 9. O'Briain gesturing making two-sided toast in (a); Navajo interpreter gesturing making toast over an open fire in (b)

Another case of intentional conceptual non-alignment with an accompanying non-aligning gesture occurring in the SR discussion is Example 6. In this case, O'Briain was explaining to the Renaissance fellows how a toilet works.

Example 6: Understanding how a toilet works

(a) (b)

Figure 10. Gesturing about using a modern toilet in (a); gesturing about flushing a "historical" pull-chain toilet in (b)

[7] Numerous researchers have discussed the interpreter's fidelity to the source or target. Nida (1964), for example uses the terms "formal equivalence" versus "dynamic equivalence". See also Gile (1995) on primary information (the content) and secondary information (background information, speaker style, etc.).

While this Spanish interpreter did not gesture pulling on a chain to flush a toilet during the interpretation, in the SR he was very clear about how he was conceptualizing the toilet in relation to the Renaissance visitors. O'Briain's point was that the visitors would not have been able to relate to a modern toilet whatsoever, and would be amazed at its efficiency in flushing away waste. He says, "You gather your robes around you. You sit down, you evacuate your waste into it, then you press a button and it's all taken away."

In the SR, however, the interpreter states the following:

> That was part of the imagery that I just described about pulling the string [here he gestures pulling the chain as in Figure 10(b)], that would be more understanding from those guys back at that time, they'd probably understand that a lot better. And so I stuck that in there to help me bridge the eras. And so, you know, that's how I figured that out, and I, I just went, I just went with it. And actually, my mind did it all by itself…. Actually, I just went with the words that he was speaking. Actually, I went straight to it, but *the image was different*. My image of it was different because I had to bridge the gap between the two eras.

The intentional non-alignment here lies in the interpreter's decision that, for whatever reason, he thought he needed to make the source text more relatable to the imaginary visitors. So, rather than adopting the image of a modern toilet that O'Briain was creating, he conceptualized a more historic version. As he states in the SR, this did not affect what he said in his interpretation – it was pointedly to assist him in working out how to fit the pieces of the text together. We thus see this as an example of a non-aligned conceptualization fully explicated by the pull-chain gesture in the SR.

However, there is an unintentional element to the non-alignment of conceptualization between the source speaker and interpreter here too; the raised tank, pull-chain toilet was invented in the 1880s, so the Renaissance folk would not know it. While we appreciate that the interpreter was consciously working through source-to-target comprehension, his conceptualization of the toilet does not in fact help the addressee—only him. In this case, we see a complex example of intentional conceptual non-alignment supported by gestural non-alignment in the SR, but with an added element of unintentional non-alignment in the conceptualization process of how to link the source text to a potential target audience.

3.3. Conceptual alignment and gestural non-alignment; conceptual non-alignment and gestural alignment

The category of conceptual alignment plus gestural non-alignment might seem illogical, because how would an interpreter align with a source speaker's conceptualization of some entity or event, and yet produce gestures that do not align with the gestures of the source speaker? Several factors suggest that this might be possible at least to some extent. First, what exactly does conceptualization entail? In the discussion above, we considered conceptualizing somewhat broadly, and while space here does not permit detailed discussion, we subscribe to the idea of conceptualization as dynamic (Langacker, 2008) and potentially complex, as the example below illustrates. There is reason to think that conceptual alignment may be partial. Second, and very much related, interpreters are trained to take on the perspective of the source speaker, to see things as they do, as exemplified in section 3.2.1 above where, in the SR, the interpreter commented "Here, I am him". However, Leeson et al. (2017a, 2017b) show that at most, interpreters have a blended viewpoint that may include some aspects of what they believe is the speaker's subjective viewpoint along with aspects clearly stemming from their own subjective view. This is illustrated in Example 7, when O'Briain talks about a fridge being a

modern appliance that would marvel his Renaissance visitors. In one Spanish interpreter's SR, the following exchange took place.

Example 7: It was *my* fridge

Researcher: We're talking about a thing – a fridge – so, what did you see?

Interpreter: I saw a fridge! (laughs)

R: Yours or not?

I: Yes! For some reason, actually … to be exact, I saw *my freezer* in my mind, which is a standup freezer.

Figure 11. O'Briain's non-specific gesture as he mentions the fridge

(a) (b)

Figure 12. The interpreter's similar one-handed gesture in the SR when she says that she saw (i.e., visualized) a fridge, in (a); a second gesture when she says "a standup freezer" in (b)

Here, when O'Briain mentions a fridge, he does not elaborate on its physical characteristics but rather begins talking about its function of keeping food cold. But he makes the gesture seen in Figure 11, which references a large thing, in a location proximal to him. We do not get to know what, if anything, he visualizes here, but after viewing the video and being asked what the interpreter saw, she says emphatically and without hesitation, "I saw a fridge", and gestures, although with just one hand (Fig. 12a), similar to O'Briain's 'thing' gesture. When asked if it was *her* fridge, she says "yes" and elaborates that she had visualized her own standup freezer. So, in this instance, there is no functional differentiation between O'Briain's and the interpreter's gesture (the fact that her version of the gesture is one-handed is immaterial), and yet what she had visualized (i.e., conceptualized) could not possibly have been what O'Briain had, and vice versa.

On one level, then, we might say that there was indeed conceptual alignment: they both conceptualized a fridge. But in fact, the interpreter's actual conceptualization was of something entirely subjective, based on salient, experiential interaction with a specific appliance, seen

via a now differentiated gesture of grasping the handle of her tall, standup freezer, shown in Figure 12(b). Thus, we suggest that this example at least in part demonstrates conceptual non-alignment, but that this non-alignment is not at first noticeable because of the similarity of the source speaker's and the interpreter's initial gestures that referred in a non-specific way to a large object. Further analysis may reveal additional examples of these two alignment categories for the spoken language participants in the study, but this is left for future examination. At least one potential example comes from one participant working from English to ASL that had to do with the wall in O'Briain's sketch. Much like O'Briain, this interpreter gestured plugging appliances into the wall, although they oriented these gestures as if the wall was directly in front of them rather than off to the side. These gestures, then, may be considered as aligned (although note Sweetser [2023] on the significance of gesture spaces), even though the conceptualization of the event is non-aligned at the abstract discourse level, where the distally positioned wall represents something unknowable and the interpreter's frontal proximal position does not reflect this (see Example 1).

4. Discussion

There are two aspects of gesture relating to the task of simultaneous interpreting. First, it is of interest whether the interpreter pays attention to the source speaker's gestures and, second, whether the speaker's gestures contribute to the interpreter's construction of speaker meaning. It has been demonstrated in both conversational and experimental data that listeners extract information from speakers' gestures not found in the speech itself (see Cooperrider & Goldin-Meadow, 2017; Hostetter, 2011; Kendon, 2004). The SRs indicate that the interpreters were not usually consciously aware of the source speakers' gestures, e.g. regarding differentiated comparative frame spaces (Example 3). One could argue that in this example the interpreter's gestures matched those of the speaker not because she took her cue from what the speaker gestured, but because gesturing toward two distinct spaces to differentiate items being compared is rather common in terms of human cognition and cognitive organization. Note the spoken construction in English that reflects this: "on one hand … and on the other …". It may not have mattered whether the speaker gestured at all in this instance because the schema of contrasting ideas occupying different gesture spaces metaphorically is available across the community of speakers. But even for an interpreter-as-speaker, these study participants' own use of such a gesture sequence appears not to be a conscious event. Example 2 (covering the world with your thumb) may be a similar example, although not in the sense of cognitive organization, but as a common experiential event like covering the moon with your thumb. Therefore, it is an accessible and meaningful action whether or not the interpreter is aware that the speaker has made this gesture (in Figure 5c the interpreter's eyes are closed during this segment; he would not have seen Hadfield make the gesture). However, Example 1 (the wall) is of a different sort, because the rightward gesture is not a general way of referring to walls. In this specific instance positioning the wall gesturally at a far-right distal location outside the normal field of vision goes well beyond a simple referential gesture to a concrete object, and rather is about the more abstract unknowability of how complex modern systems work—this is O'Briain's entire theme in his stand-up routine. In this case, whether they were aware of it or not, some of the interpreters understood this abstract meaning and the significance of the gesture, simulating the same perspectivized relationship with the wall from their own physical point of view. In doing this, they truly were enacting the subjective stance of the speaker.

This study examines how the interpreter's own gestures align with those of the source speaker as a window on conceptual alignment, and we clearly see this taking place, as in Examples 1-3. In other cases, however, this alignment breaks down (Example 4). At times, such misalignment

is not evident during the interpretation itself because revealing gestures do not appear there. However, there is an advantage to the SR exercise as a research tool in showing how the interpreter, in the SR, conceptualized various elements of the text through their gestures, as in Examples 5 (the toaster) and 6 (the toilet), whether or not this affected how they interpreted the segment into the target text. In addition to gaining information on how the study participants conceptualized aspects of the source texts, especially through visualization (see Stachowiak-Szymczak, 2019), the SR exercise gave the participants an opportunity to reflect on their work. This turned out to be an unexpected benefit for them, in that for the most part, in their day-to-day work of interpreting, the majority had not thought much of their source speakers' gestures (nor their own, for that matter). In the SR, most participants commented that paying attention to the speakers' gestures would have helped them make sense of the texts.

In the SRs, the interpreters were often surprised when they watched both the video-recorded source texts along with their interpretation of them simultaneously. Numerous times the interpreters' gestures mimicked those of the source speaker without them realizing they were doing this (see Kimbara, 2006). Interactional alignment (Feyaerts et al., 2017) involves copying behaviors in which verbal and gestural contributions by one speaker are re-used by another speaker, which results from "interactive grounding" (Feyaerts et al., 2017, p. 140), and is a factor that drives linguistic choices. During the SR one study participant commented upon seeing his gestural alignment with the source speaker's that "I guess I'm starting to mimic him, which might be a way to get into his head". This may suggest that gestural alignment can lead to conceptual alignment, not just be a reflection of it.

5. Conclusion

In discussing the role that a gesture can play in contributing meaning to a spoken utterance, Feyaerts (2023) gives the example of a Belgian politician commenting to the parliamentary assembly. She says, "This information was shared…", using a passive construction. Feyaerts' focus is on a co-speech gesture and features of the construction itself, rather than an interpretation of it. What an interpreter would have missed had they not been watching the speaker was that she gestured referentially both to herself and toward some parliamentary members, adding the very specific information that the sharing took place between particular individuals, thus disambiguating what was not specified in her spoken utterance. This example illustrates the multimodal nature of discourse that interpreters encounter and participate in. The present examination and the examples presented above explore the roles that co-speech gestures play in source speakers' discourse, the alignment of interpreters' gestures, and how these relate to conceptual alignment between the interpreter and speaker (see Kita et al.'s, 2017, gesture-for-conceptualization hypothesis, which outlines how gestures may activate, manipulate, package, and explore spatio-motoric information; Kita et al. suggest that many representational gestures are self-oriented, but can also be communicative). Stimulated Recalls (SRs) augmented the video-recorded interpretations of two source texts by fourteen study participants. While the study as a whole included both spoken and signed language interpreters, this paper reports only on data from the spoken-to-spoken language interpreters, whose working languages were English and French, Spanish, Ukrainian, or Navajo.

Facets of the interpreters' visualizations were evident in structural choices in the target texts (for example, 'cabin' in Example 4), often quite clearly in their gestures and gesture spaces during their interpretations, and in their descriptions of what they were visualizing or conceptualizing in the SRs. Comparing the source speakers' and interpreters' gestures revealed numerous instances of conceptual alignment (or non-alignment) by allowing us to see examples of adopting a speaker's viewpoint when doing so is otherwise an invisible mental

state. But it is also evident through many of these examples that the interpreters leveraged the affordances of their own subjective experiences in visualizing and conceptualizing speaker meaning. Most often, this emerged as a blended viewpoint. In "adopting" a speaker's viewpoint, the interpreter draws on their own belief as to what that (mental) viewpointed conceptualization is, without having direct access to it. This is then coupled with their own inescapable subjective conceptualization. Understanding that the interpreter's constructed comprehension of the source text necessarily results in a blended viewpoint is significant because it is a more realistic perspective of what source text comprehension is like, and how it might inform the intersubjective construction of the target text, with the target audience in mind.

6. References

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18*(5/6), 513-562. https://doi.org/10.1080/01690960344000026.

Bergen, B. (2005). Mental simulation in literal and figurative language understanding. In S. Coulson & B. Lewandowska-Tomaszczyk (Eds.), *The literal and nonliteral in language and thought* (pp. 255–278). Peter Lang.

Bloom, B. S. (1954). The thought processes of students in discussion. In S. J. French (Ed.), *Accent on teaching: Experiments in general education* (pp. 23-46). Harper.

Boogaart, R., & Reuneker, A. (2017). Intersubjectivity and grammar. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 188-205). Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*(3-4), 155-180. https://doi.org/10.1159/000261913

Chartrand, T. L., & Bargh, J.A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910. https://doi.org/10.1037/0022-3514.76.6.893

Cienki, A. (2013). Cognitive linguistics: Spoken language and gesture as expressions of conceptualization. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body – language – communication: An international handbook on multimodality in human interaction. Volume 1* (pp. 182-201). Mouton de Gruyter.

Cienki, A. (2024). Variable embodiment of stance-taking and footing in simultaneous interpreting. *Frontiers in Psychology, 15*:1429232. https://doi.org/10.3389/fpsyg.2024.1429232

Cooperrider, K., & Goldin-Meadow, S. (2017). Gesture, language, and cognition. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 118-134). Cambridge University Press.

Croft, W. (2000). *Explaining language change: An evolutionary approach.* Longman Linguistics Library, Pearson Education Ltd.

Dancygier, B. (2012). Negation, stance verbs, and intersubjectivity. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 69-93). Cambridge University Press.

Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics, 15*(2), 223-238. https://doi.org/10.1515/cogl.2004.009

Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge University Press.

Ferrara, L., & Johnston, T. (2014). Elaborating who's what: A study of constructed action and clause structure in Auslan (Australian Sign Language). *Australian Journal of Linguistics, 34*(2), 193-215. https://doi.org/10.1080/07268602.2014.887405

Feyaerts, K. (2023, August 7-11). The emergence of multimodal grammatical construct(ion)s through pointing gestures [Conference presentation]. *International Cognitive Linguistics Conference (ICLC) 16, HHU Düsseldorf, Germany*. https://iclc16.phil.hhu.de/

Feyaerts, K., Brône, G., & Oben, B. (2017). Multimodality in interaction. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 135-156). Cambridge University Press.

Frick-Horbury, D. (2002). The use of hand gestures as self-generated cues for recall of verbally associated targets. *The American Journal of Psychology, 115*(1), 1-20. https://doi.org/10.2307/1423671

Gerver, D. (1976). Empirical studies of simultaneous Interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Applications and research* (pp. 165-207). Gardener Press.

Gile, D. (1995). *Basic concepts and models for interpreter and translator training*. John Benjamins.

Hagoort, P., & Özyürek, A. (2024). Extending the architecture of language from a multimodal perspective. *Topics in Cognitive Science, 00*, 1-11. https://doi.org/10.1111/tops.12728

Hatim, B., & Mason, I. (1990). *Discourse and the translator*. Longman.

Herring, R. E., & Tiselius, E. (2020). Making the most of retrospective process tracing in dialogue interpreting research. *FITISPos International Journal, 7*(1), 53-71. https://doi.org/10.37536/FITISPos-IJ.2020.7.1.244

Hinnell, J., & Rice, S. (2016, July 18-22). 'On the one hand…': Opposition and optionality in the embodied marking of stance in North American English [Conference Presentation]. *International Society for Gesture Studies (ISGS) 7, Sorbonne Paris 3, Paris, France*. https://iclc2019.site/

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297-315. https://doi.org/10.1037/a0022128

Janzen, T. (2005). Introduction to the theory and practice of signed language interpreting. In T. Janzen (Ed.), *Topics in signed language interpreting: Theory and practice* (pp. 3-24). John Benjamins.

Janzen, T., & Shaffer, B. (2008). Intersubjectivity in interpreted interactions: The interpreter's role in co-constructing meaning. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 333-355). John Benjamins.

Janzen, T., & Shaffer, B. (2013). The interpreter's stance in intersubjective discourse. In L. Meurant, A. Sinte, M. Van Herreweghe & M. Vermeerbergen (Eds.), *Sign language research, uses and practices: Crossing views on theoretical and applied sign language linguistics* (pp. 63-84). Walter de Gruyter and Ishara Press.

Janzen, T., Shaffer, B., & Leeson, L. (2022, 11-13 November). Conceptual and gestural alignment and non-alignment in simultaneous interpreting [Conference presentation]. *High Desert Linguistic Society (HDLS) 15. University of New Mexico, Albuquerque*. https://www.unm.edu/~hdls/index.html

Janzen, T., Shaffer, B., & Leeson, L. (2023). What I know is here, what I don't know is somewhere else: Deixis and gesture spaces in American Sign Language and Irish Sign Language. In T. Janzen & B. Shaffer (Eds.), *Signed language and gesture research in cognitive linguistics* (pp. 211-242). de Gruyter Mouton.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Kimbara, I. (2006). On gestural mimicry. *Gesture, 6*(1), 36-61. https://doi.org/10.1075/gest.6.1.03kim

Kita, S. (Ed.). (2003). *Pointing: Where language, culture and cognition meet*. Lawrence Erlbaum.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245-266. http://dx.doi.org/10.1037/rev0000059

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.

Leeson, L., Shaffer, B., & Janzen, T. (2017a, 31 March – 2 April). "Let me describe sleep in space". Leveraging visual conceptualisation for interpreting practice [Conference presentation]. *International Symposium on Signed Language Interpretation and Translation Research. Gallaudet University, Washington DC*. https://gallaudet.edu/interpretation-and-translation/2017-interpretation-and-translation-research-symposium/2017-symposium-schedule/

Leeson, L., Shaffer, B., & Janzen, T. (2017b, 16-21 July). What's sleep like in space? Interpreters embodying speakers' lived experiences [Conference proceedings]. *International Pragmatics Association (IPrA) 15, Belfast, Northern Ireland*. https://cdn.ymaws.com/pragmatics.international/resource/collection/C57D1855-A3BB-40D8-A977-4732784F7B21/15th%20IPC%20abstracts-Belfast.pdf

Leonteva, A. V., Cienki, A., & Agafonova, O. V. (2023). Metaphoric gestures in simultaneous interpreting. Russian Journal of Linguistics, 27(4), 820-842. https://doi.org/10.22363/2687-0088-36189

Linell, P. (2009). *Rethinking language, mind and world dialogically: Interactional and contextual theories of human sense-making*. Information Age Publishing, Inc.

Liddell, S. K., & Metzger, M. (1998). Gesture in sign language discourse. *Journal of Pragmatics, 30*(6), 657-697. https://doi.org/10.1016/S0378-2166(98)00061-7

Neisser, U. (1976). *Cognition and reality: Principles and implications for cognitive psychology*. Freeman.

Nida, E. A. (1964). *Toward a science of translating*. E. J. Brill.

Olza, I. (2025) Modeling gestural alignment in spoken simultaneous interpreting: The role of gesture types. *Parallèles, 37*(1), 66-84. https://doi.org/10.17462/PARA.2025.01.05

Olza, I. (202). Gestural alignment in spoken simultaneous interpreting: A mixed-methods approach. *Languages, 9*(151). https://doi.org/10.3390/languages9040151

Reddy, M. (1993). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 164-201). Cambridge University Press.

Russell, D., & Winston, B. (2014). Tapping into the interpreting process: Using participant reports to inform the interpreting process in educational settings. *Translation and Interpreting, 6*(1), 102-127. https://doi.org/10.12807/ti.106201.2014.a06

Saunders, D., & Parisot, A.M. (2023). Insights on the use of narrative perspectives in signed and spoken discourse in Quebec Sign Language, American Sign Language, and Quebec French. In T. Janzen & B. Shaffer (Eds.), *Signed language and gesture research in cognitive linguistics* (pp. 243-272). de Gruyter Mouton.

Seleskovitch, D. (1978). *Interpreting for international conferences: Problems of language and communication*. Pen and Booth.

Setton, R. (1999). *Simultaneous interpretation: A cognitive-pragmatic analysis*. John Benjamins.

Stachowiak-Szymczak, K. (2019). *Eye movements and gestures in simultaneous and consecutive interpreting*. Springer.

Sweetser, E. (2023). Gestural meaning is in the body(-space) as much as in the hands. In T. Janzen & B. Shaffer (Eds.), *Signed language and gesture research in cognitive linguistics* (pp. 157-180). de Gruyter Mouton. https://doi.org/10.1515/9783110703788-007

Taylor, J. R. (2017). Lexical semantics. In B. Dancygier (Ed.), *The Cambridge handbook of cognitive linguistics* (pp. 246-261). Cambridge University Press.

Tipton, R. (2008). Interpreter neutrality and the structure/agency distinction. In C. Valero-Garcés (Ed.), *Investigación y práctica en traducción e interpretación en los servicios públicos – Desafíos y alianzas* (pp. 182-196). Universidad de Alcalá de Henares.

Tirkkonen-Condit, S., & Jääskeläinen, R. (Eds.). (2000). *Tapping and mapping the processes of translation and interpreting: Outlooks on empirical research*. John Benjamins.

Venuti, L. (1995). *The translator's invisibility: A history of translation*. Routledge.

Wadensjö, C. (1998). *Interpreting as interaction*. Longman.

Wilcox, S., & Shaffer, B. (2005). Towards a cognitive model of interpreting. In T. Janzen (Ed.), *Topics in signed language interpreting: Theory and practice* (pp. 27-50). John Benjamins.

Winston, E. A. (1995). Spatial mapping in comparative discourse frames. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 87-114). Lawrence Erlbaum.

Wu, Y. C., & Coulson, S. (2007). How iconic gestures enhance communication: An ERP study. *Brain and Language, 101*(3), 234-245. https://doi.org/10.1016/j.bandl.2006.12.003

Zagar Galvão, E. (2013). Hand gestures and speech production in the booth: Do simultaneous interpreters imitate the speaker? In C. Carapinha & I. Santos (Eds.), *Estudos de Lingüística* (pp. 115-129). Coimbra University Press.

Terry Janzen
University of Manitoba
66 Chancellors Circle
R3T 2N2 Winnipeg
Canada

terry.janzen@umanitoba.ca

Biography: Terry Janzen is a professor in the Department of Linguistics at the University of Manitoba with interests in discourse structure and the expression of viewpoint primarily in signed languages, the role of gesture in both signed and spoken language, and intersubjectivity in interpreted interactions. He edited *Topics in Signed Language Interpreting: Theory and Practice* (2005; soft cover 2008), and co-edited the volume *Signed Language and Gesture Research in Cognitive Linguistics* (2023).

Lorraine Leeson
Trinity College Dublin
House 4, Front Square
College Green
Dublin 2
Ireland

leesonl@tcd.ie

Biography: Lorraine Leeson is Professor in Deaf Studies at the Centre for Deaf Studies in the School of Linguistics, Speech and Communication Sciences, Trinity College Dublin (Ireland), where she takes a multidisciplinary approach to aspects of work on signed languages and interpreting. Lorraine is also a PI at the Science Foundation Ireland funded ADAPT Centre.

Barbara Shaffer
University of New Mexico
MSC03 2130
1 University of New Mexico
Albuquerque, NM 87131
USA

bshaffer@unm.edu

Biography: Barbara Shaffer is a professor in the Department of Linguistics at the University of New Mexico. She studies the discourse structures of American Sign Language and English with a specific focus on gesture, stance markers, and intersubjectivity. She co-edited the volume *Signed Language and Gesture Research in Cognitive Linguistics* (2023).

Parallèles 37(1) April 2025

Interpreters' gestural profiles across settings: A corpus-based analysis of healthcare, educational and police interactions

Monika Chwalczuk

University of East Anglia / Polish Academy of Sciences

Abstract

Embodied semiotic resources, such as hand gestures, are increasingly recognised as essential tools in interpreter-mediated communication. Most multimodal studies on public service interpreting (PSI) rely on microanalytic frameworks, providing fine-grained analyses of excerpts of authentic interactions. However, such case-oriented approaches are based on data sets that are too limited in size to reveal overarching patterns governing multimodal activity in PSI. This study addresses the gap by investigating dominant trends in interpreters' gesture production across healthcare, educational, and police settings. A corpus of video recordings featuring 24 interpreters is annotated in ELAN. Statistical analysis reveals minimal variation in the distribution of gesture types across contexts, with pragmatic and deictic gestures dominating. Interpreters' gestural profiles closely align with those of primary speakers, suggesting that interpreters adapt their gestural production to match speakers' multimodal activity. A qualitative analysis of 45 cases of gestural mimicry suggests that it is used as a cognitive-aid strategy, as well as a means to disambiguate lexical items and support conceptual grounding and participatory sense-making among interactants. Further research is needed to explore the cognitive mechanisms behind the recurrent patterns of interpreters' gesture production and to evaluate the impact of gestural mimicry on users' perceptions of interpreter performance.

Keywords

Public service interpreting, co-speech gestures, gestural profile, gestural landscape, mimicry, triangular mirroring

 DOI:10.17462/para.2025.01.07

1. Introducing multimodal studies in PSI

For over a decade, scholars in interpreting studies have increasingly examined interpreter-mediated interactions through the lens of *multimodality* (Pöchhacker, 2021), capturing the interplay of verbal, vocal and spatio-visual cues in creating meaning (Kress, 2010; Mondada, 2016). This framework has been particularly fruitful in *public service interpreting* (PSI), where both onsite and video-remote encounters (Davitti, 2019) showcase a range of functions supported by embodied semiotic resources, such as gestures, posture, or gaze. The goal of this paper is to contribute to this *multimodal turn* (Davitti & Pasquandrea, 2016) by examining the common gestural patterns observed in healthcare, educational and police interactions.

Most existing research investigates the functions of spatio-visual cues focusing on one specific setting. The following paragraphs overview recent studies pertaining to the fields of medical, pedagogical and judiciary interactions, to probe whether similar multimodal phenomena appear across different PSI contexts.

Doctor-patient encounters seem particularly relevant to multimodal studies, as they elicit gestures involved in pain descriptions (Rowbotham et al., 2016) and terminology-rich explanations of medical procedures (Quasinowski et al., 2023). In this vein, Gerwing and Li (2019) investigate body-oriented gestures in general practice, focusing on information transfer delivered through verbal and spatio-visual channels in interpreter's renditions. Krystallidou (2014, 2016) adopts a broader approach analysing posture, gesture and gaze in fostering patient inclusion in interpreted consultations. Gaze is also investigated from the perspective of training physicians to work with interpreters (Li et al., 2017), negotiating transition points between source speech and target renditions in medical interpreting (Pasquandrea, 2011) and optimising turn-taking in psychotherapeutic consultations (Vranjes et al., 2019, 2021).

Pedagogical interactions, typically portraying parent-teacher conferences, are studied for: the use of artefacts, such as school reports, through pointing; the influence of participants' spatial positioning on visibility and inclusion; as well as the coordination of interactions through embodied resources (Davitti, 2013; Davitti & Pasquandrea, 2017). Additionally, some multimodal accounts compare the use of gaze, gesture and body orientation in educational contexts with other PSI settings such as medical and legal interactions (Davitti, 2016, 2019; Davitti & Pasquandrea, 2013).

Interpreter-mediated police interrogations and court hearings attract studies focusing on the *spatial and visual ecology of actions* (Davitti, 2019), including spatial arrangements and object affordances. Recent studies investigate these factors in relation to visual cue access and chunking challenges in courtroom video-remote interpreting (Licoppe, & Veyrier, 2020). To date, few publications account for gestures of injured parties (Määttä & Kinnunen, 2024) and suspects (Monteoliva Garcia, 2017).

Despite differing PSI contexts and focal embodied resources, the studies reviewed share two main features: their methodology and analytical framework. First, they all use *multimodal corpora*—collections of audiovisual recordings enabling fine-grained analysis of speech and adjacent visual signals (Allwood, 2008; Knight, 2011). Second, due to the time-consuming nature of multimodal studies, the cited works predominantly rely on microanalysis of selected excerpts of larger corpora. Drawing on Conversation and Discourse Analysis (Davitti, 2019), they lead to descriptive, qualitative studies, tracing in great detail how interactions in PSI unfold. Nevertheless, despite providing invaluable insights into PSI fieldwork, such accounts present an important shortcoming. Notably, the case-study data do not enable drawing general conclusions as to recurrent patterns guiding the use of embodied resources in PSI. In other words, microanalysis shows what is *possible*, highlights what is *particularly interesting*, but

does not permit to see what is *typical*. Additionally, the variability in cues analysed (e.g. gaze, gestures, artefact manipulation) and research questions across studies makes it challenging to compare results and try to cross-examine the emerging collections of highlighted examples to build a global view of multimodal phenomena in PSI. Moreover, the reliance on short excerpts limits the ability to generate quantitative findings that indicate which of the wide range of the documented gestural patterns are the most frequently observed in interpreter-mediated interactions.

This paper aims to address this gap by (1) investigating corpora of substantial length, helping to paint a 'bigger picture' though quantitative analyses of the embodied resources used; (2) providing a contrastive multimodal analysis, focusing on the commonalities among healthcare, educational and police interactions; (3) and examining data featuring different interpreters working into various target languages within each setting to identify overarching multimodal patterns shared beyond cultural and linguistic dissimilarities. With a view to work with a sufficient amount of comparable data to support quantitative analysis, we conduct this exploratory study based on a corpus of video recordings presenting elicited interactions created for training purposes. The audiovisual materials showcase gestural production of interactants and professional interpreters performing *dialogue interpreting* (Mason, 1999) in onsite communicative events.

The novel contribution of this research lies in the fact that it reaches beyond qualitative descriptions of case-studies and attempts to provide statistical data helping to colour the white spots on the map of the use of gestures in a range of PSI settings. The added value of this approach is the possibility of determining which gestural patterns are the most frequently used, as opposed to the multimodal behaviours that might be particularly interesting for an in-depth microanalysis, but do not occur regularly. This leads to introducing the notion of interpreter's gestural profile in PSI, encompassing the dominant trends observed in interpreter-mediated interactions.

2. Focus on co-speech gestures

To ensure a common benchmark guiding the analysis of the embodied resources in PSI, we limit the scope of this research to the use of manual *co-speech gestures*, i.e. hand movements that accompany speech (for a review, see Hostetter, 2011). Gestures annotated in our corpora span representational, deictic, pragmatic, and interactive gestures, beats and emblems (see Leonteva et al., 2023 or Iriskhanova et al., 2023 for a similar choice of gesture types analysed in simultaneous interpreting).

Representational gestures convey semantic information through handshapes, movements or embodied actions illustrating referent's formal properties (Müller, 2014). *Deictic* gestures involve pointing to designate objects, locations or directions (Fricke, 2002). *Beats* denote rhythmic movements emphasising speech elements. *Pragmatic* gestures, also known as *recurrent gestures*, support stance-taking, structure discourse units, and enhance word search (Ladewig, 2014). They are recruited to facilitate parsing and fluent speech production, or express attitude towards the content of speech, rather than to convey semantic meaning (Ladewig, 2014). Given the conversational nature of our data, we coded *interactive* gestures as a separate type and following Bavelas (1992), we defined them as those performing phatic functions and coordinating turn-taking, thus regulating the interaction flow. Finally, given the inherently multicultural character of PSI corpora, we enrich the scope of annotated gestures with *emblems* – conventional, culturally-specific gestures that can replace verbal expressions like "OK" or "peace" (Matsumoto & Hwang, 2013).

3. Research questions and hypotheses

This paper investigates the use of gestures in consecutive dialogue interpreting within public service interactions, focusing on the gestures of speakers and interpreters and their similarities. We aim to answer the following research questions: Firstly, we investigate what gestural types dominate PSI interactions and if *gestural landscapes,* accounting for gesture production of all involved participants, differ across healthcare, educational and police settings. Drawing on the microanalytical studies cited above, we hypothesise that each context will present its own *gestural landscape*, with representational gestures most abundant in medical communication due to their use in pain descriptions, the dominant role of interactional and pragmatic gestures in educational settings focusing on administrative procedures, and deictic gestures most salient in police interactions given the common use of pointing towards documents and objects such as pictures or exhibits.

Secondly, we focus on interpreters' *gestural profiles* – average percentage distributions of different gesture types, calculated for all interpreters examined in a given setting. Rather than exploring individual differences among interpreters, our goal is to study if their gestural activity presents particular recurrent features typical of the *role* of the interpreter, and if their unique position in tripartite interactions leads to a distinctive multimodal performance compared to speakers. In other words, we aim to verify if the special role of the interpreters in communicative exchanges translates into a special use of gestures. We anticipate that interpreters will use more interactive gestures than other participants due to their potential for coordinating turn-taking (Davitti & Pasquandrea, 2017; Licoppe & Veyrier, 2020); and that they will show abundant production of pragmatic gestures known to play a role in facilitating speech production and word search (Ladewig, 2014).

Thirdly, the comparison between interpreters' and speakers' gesture production is also intended to shed light on instances of gestural *alignment* (Oben & Brône, 2016) or gestural *mimicry* (Kimbara, 2006) referring to copying gestures of the speakers in the interpreters' renditions. It is foreseen that interpreters will gesturally align with speakers but only when gestures are particularly salient due to the fact that a) the information conveyed in gesture is not represented in speech (i.e. deictic gestures disambiguating pronouns) or b) gesture conveys non-redundant information adding more complex meanings (i.e. representational gesture showing the size / shape of an object that is not fully described in speech; gestures illustrating motion in action descriptions).

4. Corpus collection and structure

The methodology involves analysing a multimodal corpus annotated in ELAN software (Sloetjes & Wittenburg, 2008). The recordings showcase simulated interpreter-mediated encounters prepared as training materials for future interpreters and public servants. The choice of elicited instead of naturalistic interactions was dictated by major difficulties in accessing and recording authentic PSI sessions due to ethical concerns (Davitti, 2019), as well as the aim of gathering a sufficient amount of data for statistical comparisons across settings. The raw footage totalled 4:40:51 hours, which was trimmed by removing side interactions involved in setting up encounters, commentary, and segments where participants' hands were out of frame. After such pre-processing, 68 video clips, totalling 128 minutes, were selected for analysis (see Table 1).

Settings	N Int	Languages	Original videos' duration	Selected segments' duration
Medical	8	Arabic, English, Spanish, Panjabi, Portuguese	1:30:48	34:22
Administrative	7	Bengali, English, Indonesian, Spanish	0:53:19	28:15
Police	9	Arabic, Polish, Czech, Dutch, English, French, German, Hungarian, Italian, Mandarin	2:16:44	66:06

Table 1. Features of the analysed corpora

The corpus showcases healthcare, education and police interactions, illustrating sessions in 15 languages from 8 families: Arabic, Bengali, Czech, Dutch, English, French, German, Hungarian, Indonesian, Italian, Mandarin, Panjabi, Polish, Portuguese, Spanish. All interactions used English or French as the A languages, fully comprehended by the author. As for B languages, their choice was based on accessibility of the recordings. Any reference to the utterances' semantic content in these languages is based on translation.

The audiovisual materials were sourced from open-access interpreters' training videos and recordings intended for public servants or migrant users, showcasing efficient ways of working with interpreters in Europe, the U.S., and Latin America. The videos were grouped by settings and a baseline of ecological validity was established by discarding recordings where interpreter's utterances seemed learnt by heart and recited. The remaining videos presented role-plays with genuine interpreting containing spontaneous co-speech gestures, all mediated in consecutive dialogue interpreting mode.

4.1. Healthcare settings

Medical interactions with eight interpreters included general practice, neurological, orthopaedic and surgical consultations conducted in English in combination with Arabic, Spanish, Panjabi, and Portuguese. Four conversations presented classical triangular sitting arrangements and other four occurred at a hospital patient's bedside. The participants were typically a doctor, a patient and an interpreter. Only one video additionally included a patient's family member.

4.2. Educational settings

Educational settings covered parent-teacher conferences, parents' interviews with school principals, interventions of school child welfare services, and foreign students' enrolment at university. Though children's performance and wellbeing at school were discussed in several role-plays, no minor participants were filmed. The recordings showed seven interpreters working in language combinations including Bengali, Indonesian and Spanish, always coupled with English. Typical spatial arrangements presented a parent, a teacher and an interpreter seated at a table. Two interactions showcased additional participants such as the second parent or other members of the school staff.

4.3. Police settings

Encounters with police spanned: victims' testimonies, witness statements and interrogations of suspects. Most recordings were staged at police stations, except three victims' testimonies arranged in hospitals. Besides typical three-party interactions, some included a lawyer or a clerk taking minutes. The videos illustrated work of nine interpreters working between Arabic, Polish, Czech, Dutch, German, Hungarian, Italian, or Mandarin and English or French.

5. Multimodal corpora analysis in ELAN

The corpora were manually annotated in ELAN for speech and hand gestures of all participants, yielding 5250 annotations – 3713 verbal utterances and 1537 gesture phrases. Participants' speech was segmented based on silent pauses, and a *gesture phrase* was considered as a unit of hand(s) movements containing a stroke accompanied by any other gesture phases, such as preparation, hold or recovery (Graziano & Gullberg, 2018). Even though gestures are often multifunctional, we strived to identify the primary function of each gesture phrase, resulting in assigning it to one of the categories: representational, deictic, pragmatic, interactive, beat, emblem. To test the validity of the annotation scheme, a 10% excerpt of the corpus was coded by two independent annotators according to the coding manual defining the six hand gesture types selected for the analysis. The modified Cohen's Kappa (Holle & Rein, 2013) reached 0.80, which corresponds to a substantial agreement (Landis & Koch, 1977), therefore validating the annotation scheme's reliability.

5.1. Gestural landscapes across settings

The first research question explored the distribution of gesture types across settings, using the notion of *gestural landscape* (GL) encompassing the gesture production of all participants in a given type of interaction (medical, pedagogical, judicial). Contrary to our predictions, the GL was surprisingly consistent across contexts. Pragmatic gestures occurred as the dominant type in all corpora, accounting for 34% in healthcare, 30% in education, and 42% in police interactions.

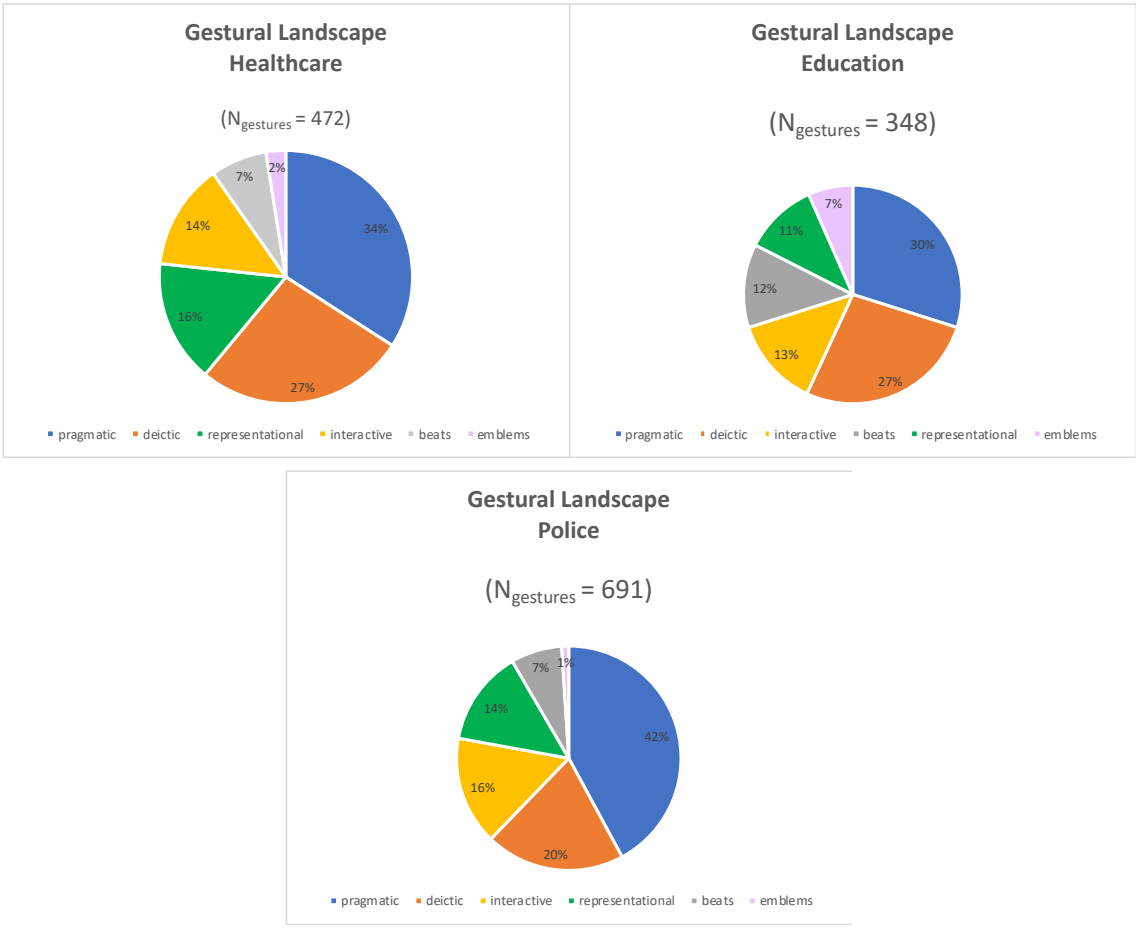


Figure 1. Gestural landscape across settings

Deictic gestures ranked second, comprising 27% in healthcare, 27%, in education, and 20% in police interactions. They served slightly different purposes depending on the context: (1) pointing towards present participants in multiparty interactions (education), (2) giving directions (education, healthcare), (3) spotlighting a discussed document, such as a report or a statement that has to be signed (education, police), (4) designating body parts in descriptions of pain or injuries (healthcare, police). Similarly to its functions documented by Vranjes and Brône (2021), pointing was recruited as a crucial semiotic resource helping to clearly distribute roles of the participants and foster housekeeping in interpreter-mediated communication.

Interactive gestures were also consistent across settings, comprising 14% in healthcare, 13% in education and 16% in police interactions. Despite their undeniable role in coordinating distribution of speech turns and visually performing floor-giving, they turned out to be far less frequent than pragmatic gestures serving internal regulation of participants' speech flow.

Representational gestures received similar scores in healthcare (16%) and police encounters (14%) but dropped down to the second-last position with 11% in educational settings. This might be explained by more imagistic content in the first two, including descriptions of accidents and pain in the first context, and recollections of physical experiences such as a robbery or an assault in the second one. Nevertheless, the percentage differences between contexts remain small, suggesting that all the examined settings present a comparable potential for iconicity.

Beats covered from 7 to 12 % of the overall gesture production, leaving them as one of the least prominent categories in the examined corpora. The exception here are pedagogical encounters where a feeble presence of imagistic content resulted in a more prominent use of gestures underscoring discourse structure (pragmatics, beats) and distribution of speech turns (interactive gestures).

Finally, to no surprise, emblems appeared as the least used type in all settings, their presence ranging from 1 to 7% of the total gesture production. Almost all their occurrences represented either different versions of a finger count or various greeting gestures, such as a formal handshake. The latter explains their increased score in educational settings in role-plays, as some video recordings spotlighted this symbolic gesture behaviour as an essential tool of establishing social relations at the beginning of interpreter-mediated interactions.

Importantly, the same tendencies were found when we excluded the interpreters' gestures from the overall gestural landscape and recalculated gesture distribution for all other participants except the interpreter (e.g. doctor and patient). Thus, the analysis showed that the effect of settings on the gestural landscape was less significant than expected. Regardless of the context, pragmatic gestures represented the lion's share of all embodied semiotic resources, followed in various combinations by deictic, interactive, and representational gestures. The settings seemed to have the most visible impact on the latter type, as healthcare consultations and police interrogations correlated with an increased use of representational gestures in comparison with educational interactions; however, the differences remained subtle. Detailed scores for each gesture type used by interactants in a given role are represented in Table 2.

Corpus	Actor		Gesture functions						
Settings	**Role**	**N[a]**	**Pragm**	**Dei**	**Rep**	**Inter**	**Beats**	**Embl**	**Total N gestures**
Healthcare	Interpreter	8[b]	64	52	20	27	13	3	179
Healthcare	Doctor	8	61	46	22	23	19	4	175
Healthcare	Patient	8	36	29	32	14	2	5	118
Education	Interpreter	7	47	44	13	19	13	3	139
Education	Teacher	7	30	39	15	22	25	12	143
Education	Parent	7	27	11	10	5	5	8	66
Police	Interpreter	9	112	60	25	35	19	4	255
Police	Police	9	33	68	10	49	20	3	183
Police	Suspect	5	34	12	16	16	1	1	80
Police	Witness	5	47	13	22	1	8	0	91
Police	Victim	7	65	11	22	7	3	0	108

Table 2. Gesture distribution across sub-corpora

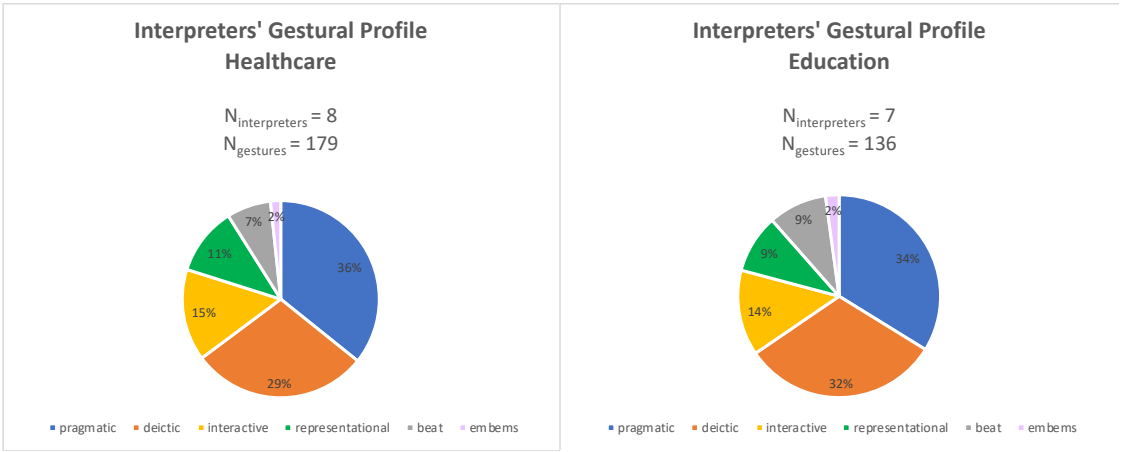
Note. Gesture types abbreviations: Pragm = pragmatic, Dei = deictic, Rep = representational, Inter = Interactive, Embl = emblem.

[a] Number of different actors observed in the same role.

[b] Data in each row represents a set of participants in the same role. For instance, the numbers describing the interpreter in healthcare settings present findings calculated on the basis of the performance of 8 different interpreters.

5.2. Interpreter's gestural profile in PSI

Next, we focused specifically on the distribution of gesture types presented in interpreters' renditions, referred to as their *gestural profile* (GP). Surprisingly, the effect of settings on interpreters' GP was minimal, with almost identical proportions of gesture types in healthcare, educational and police encounters. GPs calculated based on respectively 8, 7 and 9 interpreters' performance in each context, showed consistent distribution of 1) pragmatic, 2) deictic, 3) interactive, 4) representational, 5) beat and 6) emblem gestures, presented in order of frequency (Figure 2).



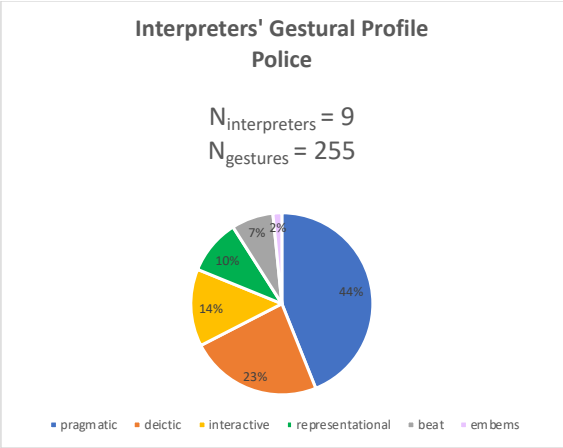


Figure 2. Interpreter's gestural profile across settings

Pragmatic gestures emerged as understandably prominent in interpreting as they support structuring the text that has to be repackaged by the interpreter in the target language. Interestingly, these gestures, used meta-communicatively to serve discursive and modal functions, appeared to be more common than interactive gestures, which are often emphasised as a key semiotic resource used by dialogue interpreters to coordinate interactions. In terms of frequency, pragmatic gestures were also far more prominent than gestures directly linked to message transfer, such as deictic and representational gestures conveying information that might not be contained in speech. The dominant presence of pragmatic gestures observed in the data corroborates findings form a study by Cienki (2024), discussing their role in performing multimodal stance-taking in interpreter's renditions. Additionally, another sub-type of pragmatic gestures observed in the data were recurrent cyclic gestures (Ladewig, 2014) that accompanied word-search, which is one of the fundamental processes involved in interpreting (Iriskhanova et al., 2023). Used as a turn-holding device, they may also support maintaining multimodal fluency by indicating that the process of searching for the right term or formulating ideas in the target language is ongoing. Specific applications of deictic and representational gestures in interpreters' renditions will be discussed in Sections 6.1–6.3.

Another pattern that emerged from a comparison between GPs of interpreters and GLs accounting for gestural activity of all participants was a strong correlation between the gesture type distribution remarked in interpreters and in all other speakers (see Figures 1 and 2). This observation grounded the next hypothesis of the study. If interpreters' profile is quite an accurate replica of the overarching gestural landscape, does it mean that interpreters systematically perform gestural alignment with the speakers?

6. Gestural alignment

To test the hypothesis regarding gestural alignment (GA), we visually inspected the corpus for instances of similar gestures produced by both speakers and interpreters. On a case-by-case basis, we determined occurrences of mimicry according to three criteria. First, paired gestures referred to (nearly) identical semantic content conveyed in speech (e.g. *to hug* and *to take somebody in one's arms*). Second, both gestures shared the same function (e.g. deictic). Third, they presented common formal features (e.g. movement direction, hand shape, tracing similar shapes, embodying similar actions). These parameters enabled reliable identification of 45 cases of mirrored representational and deictic gestures, detailed in Table 3.

Pragmatic gestures were excluded from the mirroring analysis due to their extreme variability in form (e.g. used gesture space, articulators). Since dialogue interpreters need to re-structure discourse in the second language and pragmatic gestures are recruited to support this process,

the high variability of language pairs in our data made such analysis too challenging for the scope of the present study. Furthermore, given that interpreters' pragmatic gestures can reflect stances of the source speakers or the interpreters themselves (Cienki, 2024), it would be difficult to determine which instances of similar pragmatic gestures produced by different interactants actually resulted from mirroring; this would require a more nuanced annotation scheme, too time-consuming for the size of the analysed corpus.

Settings	**Representational**	**Deictic**	**Total per settings**
Healthcare	7	10	**17**
Education	1	1	**2**
Police	12	14	**26**
Total per type	**20**	**25**	**45**

Table 3. Cases of gestural alignment in the analysed corpora

All examined corpora showed examples of gestural mimicry, nevertheless educational settings revealed only isolated cases of this phenomenon (N=2), in comparison with healthcare (N=17) and police interactions (N=26). Regarding representational gestures, this might be explained by the conversation topics generating more descriptive, imagistic content in the latter two. Consequently, contexts with a higher number of representational gestures in general correlated with more instances of their reproduction. This trend is not confirmed for deictic gestures though, as their frequent usage by both the original speakers and the interpreters in educational videos did not translate into a larger number of mirrored gestures. To shed light on the communicational contexts of engaging in gestural alignment we provide qualitative analysis of examples showcasing particular gesture types.

6.1. Disambiguation of pronouns through deictic gestures

Deictic gestures are mostly mirrored while pointing to objects and locations in a shared space. This involves bodily parts (e.g. */pain in the knee/, /my head hurts/*), documents (e.g. */please sign here/, /available on this website/*) or directions (e.g. */lay down here, with your head up here/*). Much like in monolingual settings, they help to map speech content onto physical referents, following the principle of contiguity (Fricke, 2002). In PSI, it is however the use of deictic gestures for designating people that seems an essential asset to the interpreter, as they aid disambiguation of personal pronouns (e.g. you, your, yours). These might become sources of confusion in PSI, as professional interpreters are trained to use the first person to recreate speakers' utterances, and the third person singular to refer to themselves as interpreters. This laminated nature of interpreter's utterances (Vranjes & Brône, 2021) becomes apparent in repairs such as */excuse me doctor/ /the interpreter does not understand one of the terms that [the patient] she's using/ /and I would like to ask for clarification/* performed with a deictic gesture towards the interpreter accompanying the first-person pronoun 'I'.

Moreover, the direction of mirrored pointing gestures is adjusted to refer to same *participant* but not the same *speaker.* For instance, Figure 3[1] shows deictic gestures respectively used by the doctor while introducing the interpreter, and by the interpreter herself when rendering the same utterance.

[1] Still images stem from an educational video produced by dr Charles Liao at the Stanford School of Medicine. The materials are available on YouTube: https://www.youtube.com/watch?v=Uhzcl2JDi48. Written consent of the copyright holders was obtained to use the screenshots in the present paper.

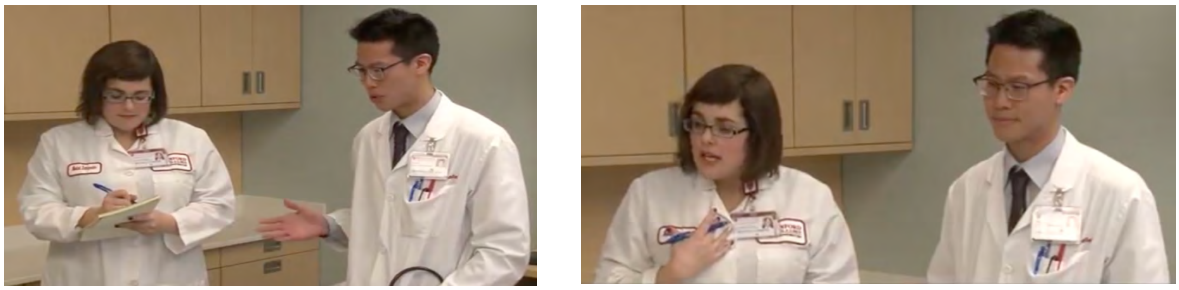


Figure 3. Disambiguation of personal pronouns using deictic gestures

In both cases gestures refer to the interpreter, even though they are produced by different speakers. In this particular example, since the interpreter is wearing a gown, she could easily be mistaken for medical staff. The use of visual resources clarifies participants' roles, and helps to structure interpreter's utterances that can embody discrepant voices of different people (cf. Vranjes & Brône, 2021).

6.2. Mirroring representational gestures – reflection of embodied cognition

As hypothesised, recurrent contexts of mirroring representational gestures mostly involved: introducing unfamiliar terminology, such as medical procedures in doctor's consultations, or referring to physical experiences, especially pain or violence in police and healthcare settings. Apart from typical dyadic mimicking of a speaker's gesture by an interpreter, we documented instances where this scheme was extended to several gestures or multiple participants. Figure 4[2], presenting screenshots selected from a simulated police interaction in German-French, displays a case of reproduction involving a short sequence composed of multiple representational gestures accompanying a description of an assault. The upper-panel pictures present victim's gestures recalling actions of the attacker: */he grabbed the hood of my coat/ / he turned me around/ /and he just slapped me in the face/*.

Each picture stands for one of the key actions named by the speaker, [GRAB], [TURN AROUND], [SLAP]. The lower panel shows the interpreter recreating these gestures with minimum delay, as if she was embodying the actions of the speaker. It is worth mentioning that in this moment, the source speech accelerated causing the interpreter to drop note-taking and switch to quasi-simultaneous renditions to cope with an uninterrupted information flow coming from the victim. This change of dynamics might have resulted in an instinctive adjustment of the interpreter's memory-supporting strategy. Spatio-visual information that could have been previously stored in notation as symbols or drawings needed to be allocated to a different pool of semiotic resources, hence its embodied representation, possibly supporting the lexical retrieval of the target words (Leonteva et al. 2023, cf. Morsella & Krauss, 2004).

[2] Still images stem from an educational video produced by ESIT within the framework of the IMPLI project (*Improving Police and Legal Interpreting*). The materials are available on the website of the project coordinated by the University of Bologna: https://site.unibo.it/interpretazione-giuridica-impli/en/educational-videos. Written consent of the copyright holders was obtained to use the screenshots in the present paper.

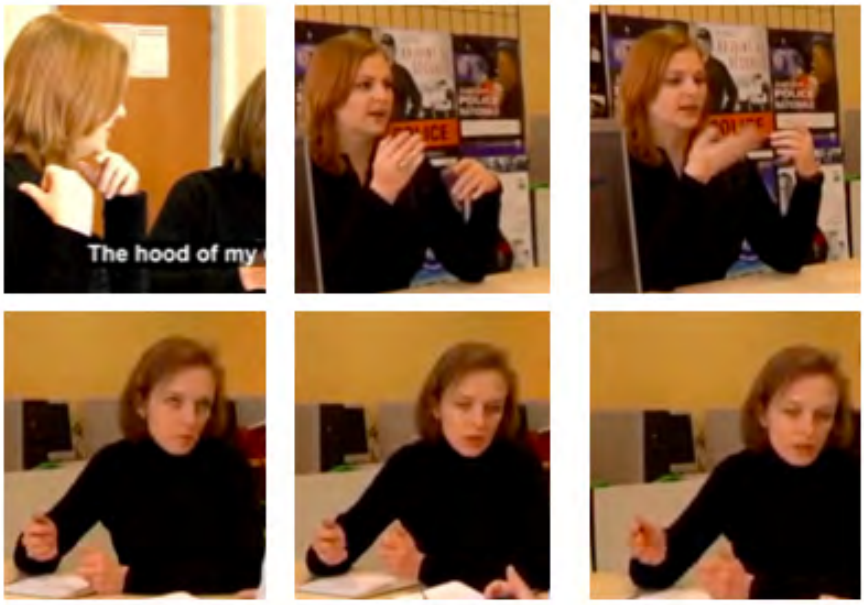


Figure 4. Reproduction of a sequence of representational gestures

6.3. Triangular gesture mirroring in participatory sense-making

Further analysis revealed that gestural alignment can be extended beyond the nuclear speaker-interpreter pair. The data contained cases of *triangular mirroring* where a key gesture introduced by a participant was mimicked not only by the interpreter, but also incorporated into user's responses. The term *triangular mirroring* is meant to capture the gesture spreading network going beyond the basic initiator-imitator pair and involving at least three different interlocutors recycling a similar gesture referring to the same semantic content. The sequence can be initiated by either leader roles (e.g. doctor) or follower roles (e.g. patient), though no interpreter-initiated patterns were found.

In most cases, triangular mirroring involved deictic gestures referring to (1) participants in the interactions, especially if their role needed to be explained (e.g. lawyer), (2) objects that the lead speakers wanted to focus attention on (e.g. a knife, a photograph presented as evidence during a police interrogation), (3) or locations in the same room (e.g. the couch where the patient was to be examined). Such use of gestural alignment confirms its role in establishing common ground by mapping the content of speech onto the physical environment (Barsalou, 2008; Beinborn et al., 2018).

As for representational gestures, triangular mirroring mostly occurred in the medical field where it accompanied introducing new concepts (e.g. terminology of medical procedures) and/or challenging explanations (e.g. pain descriptions). The latter is presented in Figure 5[3] where gestures of the patient illustrating her stomach-ache were picked up on by the doctor who used them to solicitate a more accurate description of the type of pain: [CRAMPY], as opposed to a constant pain, represented by a different iconic gesture.

[3] Still images come from the educational video produced by the University of Nottingham, available on YouTube: https://www.youtube.com/watch?v=N8iqH9qwIAQ.

Figure 5. Triangular mirroring: gesture [CRAMPY PAIN]

Embodiment appears here as an essential technique of visually illustrating physical experience that lay patients might struggle to name (Rowbotham *et al*., 2016). In discordant language communication, the handshape and intensity of representational gestures not only offer diagnostic clues, but also help to pass bits of information directly between the doctor and the patient. Gerwing and Li (2019, p. 177) report that 70% of body-oriented gestures produced by doctors and patients convey information not included in speech. Furthermore, Rivera Baldassari (2024) points out that migrant patients often struggle to accurately describe pain, even in their mother tongues, leading to additional challenges for interpreting. From this point of view, representational gestures grant a more reliable means of communicating bodily experiences, as they help to by-pass the linguistic and terminological barrier by adding embodied representations that clarify the meaning of new and unclear concepts. Reproduction of such gestures by the interpreter reassures the interactants (Gerwing & Li, 2019), confirming that their contributions to the *participatory meaning-making process* (De Jeagher & Di Paolo, 2007) have been received correctly, thus creating of a shared repertoire of word-gesture entities.

7. Discussion

The findings suggest that the variation in the use of co-speech gestures across healthcare, educational and police PSI interactions is less pronounced than expected. Firstly, the comparison of gestural landscapes encompassing all participants reveals similar frequency distributions, characterised by a dominant use of pragmatic and deictic gestures. The latter appear as an inherent tool for establishing multimodal mappings connecting abstract linguistic items with referents present in a shared interactional space, be there objects, places or people. The crucial role of pointing in language-discordant communication is confirmed through its presence in gestural mimicry and triangular mirroring, underscoring its usefulness both for primary speakers and interpreters. Thus, our quantitative data corroborates findings from earlier microanalytical studies examining medical (Gerwing and Li, 2019), pedagogical (Davitti & Pasquandrea, 2017) and police interactions (Monteoliva Garcia, 2017).

The effect of settings is most noticeable in representational gestures, which cover a larger percentage of the gestural landscape in healthcare and police interactions in comparison with educational settings, though the differences remain subtle. The analysis reveals that even though each of the examined settings involves its own conversation topics and communicative goals, the overall gestural landscape presents far more similarities than differences across settings.

Secondly, zooming in on the gestural profiles of the interpreters, we observe that they remain consistent regardless of interactional settings, suggesting that the universal challenges of onsite

dialogue interpreting have a stronger impact on the interpreter's gesture production than local difficulties related to particular conversation topics. We note that the overall gestural profile of the public service interpreter is composed mainly of pragmatic and deictic gestures, followed by interactive and referential ones. The predominant role of pragmatic gestures, outnumbering any other kind, is consistent with findings from simultaneous interpreting (Iriskhanova et al., 2023) where this gesture type prevailed both in salient and non-salient occurrences; hence we observe that their position among other co-speech gestures is insensitive not only to settings, but also to modes of interpreting.

Furthermore, comparisons of gestures used by primary speakers and interpreters reveal only minor dissimilarities. This indicates that interpreters' renditions follow very similar patterns to those characterising spontaneous speech productions by other participants, hence the special role of the interpreter does not necessarily lead to a distinct use of gestures. Though certain gesture types might be particularly helpful in resolving turn-taking or disambiguation difficulties in interpreting (Vranjes & Brône, 2021), they are not applied frequently enough to shift the proportions of gestures composing interpreters' GP.

Similarities between the interpreters' gestural profiles and the overall gestural landscapes support the view that the way interpreters gesture is highly influenced by the gestural production of primary speakers. Nevertheless, though instances of gestural alignment have indeed been identified in all the settings examined here, they merely accounted for a small portion of all interpreters' gestures. This finding is consistent with Gerwing and Li (2019, p. 174) reporting that in medical encounters only 42% of speaker's body-oriented gestures were incorporated in interpreters' renditions.

Qualitative analysis of gestural alignment cases suggests that mimicry involving representational gestures seems to be deployed as a cognitive-aid strategy when interpreters deal with rich information units related to speakers' bodily experience. A possible explanation is that switching from an embodied representation of an action in the original utterance to its purely verbal description it the target speech would create additional cognitive load resulting from passing from one modality to another. Since interpreters are known to work on the verge of exhausting their mental processing space while juggling with two languages (Gile, 1995; Seeber & Arbona, 2020), it is plausible that maintaining information within one modality helps to regulate instant cognitive effort. Tapping into bodily experience seems an efficient way of connecting embodied meanings across languages, by virtue of activating the process of *conceptual grounding* (Barsalou, 2008) or *multimodal grounding* (Beinborn et al., 2018). Both terms convey the idea that *language is grounded in perceptual experience and sensorimotor interactions with the environment* (Beinborn et al., 2018, p. 2326). Its reflection on the human cognition is that during language comprehension and production the brain simulates perceptual and motor activities associated with the described situation. For instance, neuroimaging studies show that hearing action verbs associated with movement provokes activity in the motor cortex as if the hearer was performing the action themselves (Beinborn et al., 2018; Garagnani & Pulvermüller, 2016; Marstaller & Burianová, 2014). Findings from neurocognitive research help to explain how mirroring representational gestures may serve self-regulating functions for the interpreters themselves, grounding their processes of comprehension of the source and production of the target message in embodied resources that are not encoded in any language but rooted in sensorimotor experience (cf. Janzen et al., *this volume*). In a similar vein, gesture studies report that the use of representational gestures supports lexical retrieval (Morsella & Krauss, 2004), accompanies fluent speech production (Graziano & Gullberg, 2018) and facilitates speech production in L2 (Morett et al., 2012). Hence, given strong evidence that gesturing enhances a number of cognitive-linguistic operations involved in spontaneous

speech, there is no reason why their facilitatory functions would not extend to interpreters' renditions. Moreover, alignment and triangular mirroring of deictic gestures emerge as effective means of disambiguating personal pronouns referring to interpreters and speakers. Hence, our data confirms that pointing helps to maintain clarity given the laminated nature of interpreters embodying different voices (Vranjes & Brône, 2021).

Finally, from an interactional perspective, gestural mimicry contributes to multimodal *participatory sense-making* (De Jeagher & Di Paolo, 2007), where visible and spoken components of speakers' utterances are coordinated in the process of creating and negotiating meaning. Thanks to embodied resources, users who do not master the host country's language can share partial information directly with civil servants, bypassing the delay of interpreting. These observations align with Chwalczuk (2021, p. 356) who documented cases where in child psychiatry users mimicked therapist's gestures as part of backchannelling, performed without producing full speech turns, or signalling self-selecting for floor-taking as a way to manifest understanding of the source speech before interpreting was delivered. Such cases present an important alteration to the typical interpreter-mediated interactional dynamics as the interpreter is temporarily omitted in the communicational chain (Gerwing & Li, 2019). Regarding patients with limited language proficiency, visual access to key information units can prompt them to seek more active and independent participation in the interactions, performed through *mimicry*, that is reported to provide (Kimbara, 2006, p. 59):

> a resource for organizing co-participation (…) through the display of the shared form–meaning mapping, and thus, by making one's own representational action in coordination with the other's. Once associated with meaning, the form of a speaker's gesture, together with the meaning, is added to the common ground. That is, the unit of gesture and speech becomes shared by the speakers.

From the point of view of public service interpreting, the elaboration of these shared gesture-speech units may help migrants to gain a sense of agency, as embodied cues grant a possibility to communicate, even to a limited extent, directly with the public servant who represents the host country (Chwalczuk, 2021; Gerwing & Li, 2019). Moreover, sociolinguistic research on mimicry shows that it facilitates mutual understanding, fosters bonding and enhances empathy among participants, overall making interaction feel smoother (Stel & Vonk, 2009). Thus, gestural mimicry involving migrant participants may be viewed as a first step to becoming not only increasingly independent users of the dominant language (Morett et al., 2012), but also more socially integrated participants in the system of public institutions. Consequently, gestural alignment appears to play not only self-regulatory functions, activating interpreter's resources of embodied cognition, but also interactive functions fostering social inclusion and increased participation of migrants.

8. Conclusions, limitations and future directions

The findings suggest that gestural activity of the interpreter presents similar multimodal patterns across different settings. However, an obvious limitation to the study is that it uses elicited data and does not rely on authentic interpreter-mediated events. As PSI typically features sensitive and ethically challenging content, the access to ecologically valid video recordings remains to date highly restricted. Moreover, given laborious, time-consuming manual annotation of the visual content, it is only natural that corpus-based research of multiparty interactions remains a trade-off between the quantity of analysed material and the depth of the analysis. In this vein, we recognize that the corpora investigated in this paper presented a great variability of languages and geographical-cultural contexts, which may be perceived both as a weakness

or strength of the study. On the one hand, the number of videos within the same language pairs were not large enough to produce any general conclusions about the use of gestures in interpreting for/from a particular culture. However, one may claim that if despite such variability of cultural contexts, we still observe strong overarching multimodal patterns, they are more likely to be generalisable to all PSI.

The corpus offers insights in the workings of embodied cognition and conceptual grounding. However, even though the present study enables us to confirm the general tendency for gestural mimicry documented in all examined PSI settings, not only does it not explain most interpreter's gestures, but also, it does not allow to determine *what* stimulates this phenomenon in interpreting and *how it affects* the communicative efficiency of interpreter-mediated interactions. Hence, more experimental studies are required to disentangle the causes and effects of interpreter's gestural production in PSI. To this end, psychophysiological studies are needed to further investigate the links between the use of gestures in interpreting and the underlying cognitive operations, looking at the process in controlled, laboratory settings to separate self-regulatory gestural activity from its social covariates. In parallel, given that gestural alignment is observed not only in authentic, but also role-played interactions without genuine social stakes, it would be enriching to measure the impact of the interpreter's gestural style on how users of the interpreting services perceive the interpreter's professional performance depending on the gestural profile they exhibit.

All in all, multimodal corpora analyses offer valuable insights in how embodied semiotic resources are recruited in interpreter-mediated interactions. Findings of such studies ascribed to the *multimodal turn* should find their way to interpreting pedagogy and training of public servants involved in PSI interactions, as they help to shift the focus from the language barrier to a shared point of reference: the physical experience of human bodies.

9. Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Maria Sklodowska-Curie grant agreement No 847639. The CoGCIn Project—Cognitive Processes Behind the Use of Gestures in Consecutive Dialogue Interpreting—was conducted by Dr. Monika Chwalczuk as the Principal Investigator and implemented at the Institute of Psychology, Polish Academy of Sciences.

10. References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*(4), 469–489.

Beinborn, L., Botschen, T., & Gurevych, I. (2018). Multimodal grounding for language processing. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 2325–2339). Santa Fe, New Mexico, USA, August 20-26, 2018.

Cienki A (2024) Variable embodiment of stance-taking and footing in simultaneous interpreting. *Frontiers in Psychology, 15*, 1429232. https://doi.org/10.3389/fpsyg.2024.1429232

Cienki, A., & Iriskhanova, O. K. (2020). Patterns of multimodal behavior under cognitive load: An analysis of simultaneous interpretation from L2 to L1. *Voprosy Kognitivnoy Lingvistiki*, *1*, 5–11.

Chwalczuk, M. (2021). Co-speech gestures in public service interpreting: contextual analysis of multimodal corpora [Doctoral dissertation, Université Paris Cité]. https://www.theses.fr/2021UNIP7020

Davitti, E. (2019). Methodological explorations of interpreter-mediated interaction: Novel insights from multimodal analysis. *Qualitative Research*, *19*(1), 7–29. https://doi.org/10.1177/1468794118761492

Davitti, E., & Pasquandrea, S. (2017). Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings. *Journal of Pragmatics*, *107*, 105–128. https://doi.org/10.1016/j.pragma.2016.04.008

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, *6*(4), 485–507. https://doi.org/10.1007/s11097-007-9076-9

Fricke, E. (2002). Origo, pointing, and speech: The impact of co-speech gestures on linguistic deixis theory. *Gesture*, *2*(2), 207–226. https://doi.org/10.1075/gest.2.2.05fri

Garagnani, M., & Pulvermüller, F. (2016). Conceptual grounding of language in action and perception: A neurocomputational model of the emergence of category specificity and semantic hubs. *European Journal of Neuroscience*, *43*(6), 721–737. https://doi.org/10.1111/ejn.13145

Gerwing, J., & Li, S. (2019). Body-oriented gestures as a practitioner's window into interpreted communication. *Social Science & Medicine*, *233*, 171–180. https://doi.org/10.1016/j.socscimed.2019.05.040

Gile, D. (1995). *Basic concepts and models for interpreter and translator training.* John Benjamins Publishing Company.

Graziano, M., & Gullberg, M. (2018). When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. *Frontiers in Psychology*, *9*, 879. https://doi.org/10.3389/fpsyg.2018.00879

Holle, H., & Rein, R. (2013). The modified Cohen's Kappa: calculating interrater agreement for segmentation and annotation. In H. Lausberg (Ed.), *Understanding body movement.* (pp. 261–278). Peter Lang.

Hostetter, A. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*, 297–315.

Iriskhanova, O. K., Cienki, A., Tomskaya, M. V., & Nikolayeva, A. I. (2023). Silent, but salient: Gestures in simultaneous interpreting, *Research Result: Theoretical and Applied Linguistics*, *9*(1), 99–114. https://doi.org/10.18413/2313-8912-2023-9-1-0-7

Kimbara, I. (2006). On gestural mimicry. *Gesture*, *6*(1), 39–61. https://doi.org/10.1075/gest.6.1.03kim

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Krystallidou, D. (2014). Gaze and body orientation as an apparatus for patient inclusion into/exclusion from a patient-centred framework of communication. *The Interpreter and Translator Trainer*, *8*(3), 399–417. https://doi.org/10.1080/1750399X.2014.972332

Krystallidou, D. (2016). Investigating the interpreter's role(s): The A.R.T. framework. Interpreting. *An International Journal of Research and Practice in Interpreting*, *18*(2), 172–197. https://doi.org/10.1075/intp.18.2.04kry

Ladewig, S. H. (2014). Recurrent gestures. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & J. Bressem (Eds.), *Body–language–communication: An international handbook on multimodality in human interaction* (Vol. 2, pp. 1558–1575). De Gruyter Mouton.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Leonteva, A. V., Cienki, A., & Agafonova, O. V. (2023). Metaphoric gestures in simultaneous interpreting. *Russian Journal of Linguistics*, *27*(4), 820–842. https://doi.org/10.22363/2687-0088-36189

Li, S., Gerwing, J., Krystallidou, D., Rowlands, A., Cox, A., & Pype, P. (2017). Interaction—A missing piece of the jigsaw in interpreter-mediated medical consultation models. *Patient Education and Counseling*, *100*(9), 1769–1771. https://doi.org/10.1016/j.pec.2017.04.021

Licoppe, C., & Veyrier, C.-A. (2020). The interpreter as a sequential coordinator in courtroom interaction: 'Chunking' and the management of turn shifts in extended answers in consecutively interpreted asylum hearings with remote participants. *Interpreting*, *22*(1), 56–86. https://doi.org/10.1075/intp.00039.lic

Määttä, S., & Kinnunen, T. (2024). The interplay between linguistic and non-verbal communication in an interpreter-mediated main hearing of a victim's testimony. *Multilingua*, *43*(3), 299-330. https://doi.org/10.1515/multi-2023-0153

Marstaller, L., & Burianová, H. (2014). The multisensory perception of co speech gestures – A review and meta-analysis of neuroimaging studies. *Journal of Neurolinguistics*, *30*, 69–77 https://doi.org/10.1016/j.jneuroling.2014.04.003

Matsumoto, D., & Hwang, H.-S. (2013). Emblematic Gestures (Emblems). *The Encyclopedia of Cross-Cultural Psychology*, 464–466. https://doi.org/10.1002/9781118339893

Mason, I. (1999). *Dialogue interpreting*. St Jerome Publishing.

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.12177

Monteoliva Garcia, E. (2017). Police interviews with suspects: The collaborative construction of the stand-by mode of interpreting in police interviews with suspects [Doctoral dissertation, Heriot-Watt University]. https://hdl.handle.net/10399/123456

Morett, L. M., Gibbs, R. W., & MacWhinney, B. (2012). The role of gesture in second language learning: Communication, acquisition, & retention. *Cognitive Science*, *34*(8), 1457–1480. https://doi.org/10.1111/j.1551-6709.2012.01256.x

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology*, *117*(3), 411–424. https://doi.org/10.2307/4149008

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & J. Bressem (Eds.), *Body–language–communication* (pp. 1687–1702). De Gruyter Mouton.

Oben, B., & Brône, G. (2016). Explaining interactive alignment: A multimodal and multifactorial account. *Journal of Pragmatics*, *104*, 32–51. https://doi.org/10.1016/j.pragma.2016.07.002

Pasquandrea, S. (2011). Managing multiple actions through multimodality: Doctors' involvement in interpreter-mediated interactions. *Language in Society*, *40*(4), 455–481. https://doi.org/10.1017/S0047404511000479

Pöchhacker, F. (2021). Multimodality in interpreting. In Y. Gambier & L. Van Doorslaer (Eds.), *Handbook of Translation Studies* (5th ed., pp. 152–158). John Benjamins. https://doi.org/10.1075/hts.5.mul3

Quasinowski, B., Assa, S., Bachmann, C., Chen, W., Elcin, M., Kamisli, C., … Wietasch, G. (2023). Hearts in their hands—Physicians' gestures embodying shared professional knowledge around the world. *Sociology of Health & Illness*, 45(5), 1101–1122. https://doi.org/10.1111/1467-9566.13639

Rivera Baldassari, G. (2024), Sobre el dolor y otros demonios [Conference presentation]. III Congreso Internacional de Traducción e Interpretación, Lima, Peru. https://custom.cvent.com/.pdf

Rowbotham, S.J., Holler, J., Wearden, A., & Lloyd, D.M., 2016. I see how you feel: Recipients obtain additional information from speakers' gestures about pain. Patient interplay of speech and co-speech hand gestures in the description of pain sensations. *Speech Communication*. *57*, 244–256. https://doi.org/10.1016/j.specom.2013.04.002.

Seeber, K. G., & Arbona, E. (2020). What's load got to do with it? A cognitive-ergonomic training model of simultaneous interpreting. *The Interpreter and Translator Trainer*, *14*(4), 369–385. https://doi.org/10.1080/1750399x.2020.1839996

Sloetjes, H., & Wittenburg, P. (2008). Annotation by Category – ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008) Marrakech, Morocco* (pp. 816–820). European Language Resources Association (ELRA).

Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology*, *101*(Pt 2), 311–323. https://doi.org/10.1348/000712609X465424

Vranjes, J., & Bot, H. (2021). Optimizing turn-taking in interpreter-mediated therapy: On the importance of the interpreter's speaking space. *The International Journal of Translation and Interpreting Research*, *13*(1). https://doi.org/10.12807/ti.113201.2021.a06

Vranjes, J., Bot, H., Feyaerts, K., & Brône, G. (2019). Affiliation in interpreter-mediated therapeutic talk: On the relationship between gaze and head nods. *Interpreting*, *21*(2), 220–244.

Vranjes, J., & Brône, G. (2021). Interpreters as laminated speakers: Gaze and gesture as interpersonal deixis in consecutive dialogue interpreting. *Journal of Pragmatics, 181*, 83–99. https://doi.org/10.1016/j.pragma.2021.05.008

Monika Chwalczuk
University of East Anglia / Polish Academy of Sciences
University of East Anglia, Research Park
Norwich NR4 7TJ
United Kingdom

m.chwalczuk@uea.ac.uk

Biography: Dr. Monika Chwalczuk is a Lecturer in Translation and Interpreting Studies at the University of East Anglia. She previously worked as an Assistant Professor at the Psycholinguistics and Cognitive Psychology Lab at the Polish Academy of Sciences. Her project supported by a Marie Curie Research Fellowship (2022–2024) combined the fields of cognitive interpreting studies and gesture studies. From 2017 to 2022, she lectured at Université de Paris (France) and the University of Warwick (UK). She holds a PhD in Translation Studies from Université Paris Cité, where she investigated co-speech gestures in public service interpreting through the lens of multimodal corpora.

Parallèles 37(1) April 2025

Exploring the semiotic complexity of the information exchange process in healthcare interpreting: How gestural omissions and additions can impact the amount and type of information exchanged

Inez Beukeleers[1], Laura Theys[1], Heidi Salaets[1], Cornelia Wermuth[1], Barbara Schouten[2], Geert Brône[1]

[1]KU Leuven, Faculty of arts, [2]University of Amsterdam, Faculty of Social and Behavioural Sciences

Abstract

A large body of research shows that interpreters actively shape meaning and can make changes to the originals in order to coordinate mutual understanding. In this paper, we broaden the discussion by investigating the potential impact of gestural shifts on the information exchange process and the coordination of common ground in interpreter-mediated medical encounters. A qualitative analysis of three excerpts shows that omitting and/or adding representational iconic and deictic gestures can potentially lead to changes in meaning, i.e., less/more concrete renditions. Moreover, as visualization is considered a cognitive aid strategy, omitting or adding gestures can make it more/less demanding for patients to capture the full meaning of the rendered composite utterances. However, the gestural shifts can, but may not necessarily, lead to communicative troubles. This paper thus supports the idea that interpreting entails an act of strategic decision-making, yet stresses the importance 1) of raising awareness about the use of the gestures amongst interpreters and 2) of informing healthcare providers about the complexity of integrating visual information in dialogue interpreting. This paper is therefore also a warm invitation to both parties to collaboratively seek for effective strategies to cope with the semiotic complexity of healthcare communication.

Keywords

Information exchange process, coordinating mutual understanding, gestural shifts, multimodality, healthcare interpreting

 DOI:10.17462/para.2025.01.08

1. Introduction

A topic that has received a significant amount of scholarly attention in the field of Interpreting Studies is the relation between the primary participants' utterances, i.e., the originals, and the interpreters' renditions in terms of accuracy, completeness and fidelity (Pöchhacker, 2022). In order to evaluate the interpreters' performances, many researchers focus on the interpretation product and quantify the occurrences of deviations or so-called errors (e.g., Aranguri et al., 2006; Barik, 1992; Flores et al., 2003). Wadensjö (1998), however, rather opts for a more descriptive approach to the study of "originals" and "renditions" and argues that interpreting entails strategic decision-making and thus that interpreters can opt to modify, omit or add information in order to accomplish message equivalence. In that way, altering or omitting (parts) of the originals might even be recommendable in order to achieve accuracy within a particular interactional context (Wadensjö, 1998; see also Cirillo, 2012; Major & Napier, 2012).

In this paper, we aim to explore how gestural shifts in the interpretation process can potentially impact the information exchange process and the coordination of mutual understanding in healthcare interpreting. Through a qualitative analysis of three excerpts taken from authentic interpreter-mediated medical consultations, we investigate how gestural omissions and additions can potentially lead to shifts in meaning between the primary participants' utterances and the interpreter's renditions. In doing so, we thus explore how particular types of gestural shifts in the interpreting process can impact the negotiation of meaning and the coordination of common ground.

The reasons for analyzing multimodal shifts, i.e., gestural omissions and additions, in this type of discourse are twofold. First, successful information exchange is a key factor in healthcare communication (e.g., De Haes & Bensing, 2009; Menichetti et al., 2021). Clear and viable information is essential to achieve successful healthcare and good health outcomes, and to promote patient participation and shared decision-making. Yet, a large number of studies indicate that exchanging medical information is especially challenging in language discordant medical consultations (e.g., De Wilde et al., 2019; Jacobs et al. 2017). Second, existing studies already point towards the importance of the use of gestures and other visual resources in this specific setting. Healthcare providers (henceforth HCPs) use, for instance, a variety of cognitive aid strategies, including repetitions and simplifications, but also visual information such as pictures and drawings (Riloff et al., 2014; Menichetti et al., 2021) and iconic gestures that depict, for instance, medical procedures (Beukeleers et al., 2023). Therefore, we believe that interpreter-mediated medical consultations provide us with a good empirical testbed to investigate multimodal interpreting strategies.

In the following sections, we first provide a brief overview of existing studies on the relation between originals and renditions and the impact of shifts in the interpreting process in healthcare settings (Sections 1.1 and 1.2). In section 1.3, we zoom in on insights derived from Gesture Studies and Cognitive Linguistics that elaborate on multimodal meaning construction and support the idea that gestural shifts, i.e., adding, modifying and/or omitting gestures, in the interpretation process can sometimes lead to shifts in meaning (Section 1.3). Subsequently, we introduce the aims of the current study (section 2) and elaborate on the methodology (section 3). Section 4 then presents 3 excerpts that illustrate how gestural shifts in the interpreters' rendition can impact the information the patient receives. Finally, in section 5, we discuss some implications of our analyses on the conceptualization of the information exchange process in Interpreting Studies and on both HCPs' and interpreters' communicative practices.

1.1. On the relation between originals and renditions in healthcare interpreting

As exchanging clear and accurate information is essential in medical encounters, many scholars have investigated the relation between the utterances of the primary participants, i.e., originals, and the interpreters' renditions thereof in healthcare interpreting. When reviewing the literature, it appears that there are two approaches to studying this. The first one entails researchers comparing the originals with the interpreters' renditions, and quantifying and categorizing interpretation errors, often in order to evaluate interpreters' performances (e.g., Aranguri et al., 2006; Flores et al., 2003; Hsieh, 2016). Overall, these studies indicate that interpreting errors, especially omissions, are omnipresent. Moreover, most of the interpreting errors—especially those made by informal interpreters—had a potential clinical impact, for instance because the interpreter omitted information related to the dose, frequency and duration of a particular medicine (Flores et al., 2003). Therefore, authors adopting this point of view often stress that "faithful transmissions" of all utterances should be the main focus of interpreting training programs (Flores et al., 2003, p. 10).

The second approach to the analysis of the relation between originals and renditions in healthcare interpreting does not merely describe shifts in the interpreting process as "good" or "bad" in terms of the quality or "faithfulness" of the translation, but rather starts from the idea that interpreting is a situated practice and an act of strategic decision-making (e.g., Angelelli, 2004, 2019, Major & Napier, 2012; Wadensjö, 1998). Scholars working within this framework aim to capture—and thus describe, rather than prescribe—the different roles that interpreters adopt. In doing so, they describe the different interpreting strategies and their impact on the interaction.

Overall, these studies show that interpreters modify and reshape the primary participants' utterances, i.e., they omit, reduce and expand the originals. On the one hand, zooming in on elements that are often omitted, it appears that interpreters often leave out cohesive elements, such as conjunctions (Major & Napier, 2012) and affective elements, such as emotions or empathic responses to emotions (e.g., Amato, 2004; Bolden, 2000; Cirillo, 2012; Davidson, 2000; Gutierrez et al., 2019; Major & Napier, 2012; Theys et al., 2023). On the other hand, interpreters also expand the primary participants' utterances in their interpretation by, for instance, making implicit information more explicit (Major & Napier, 2012; Theys et al., 2023), adding repetitions and/or adding cohesive elements (Major & Napier, 2012). Moreover, interpreters even add zero renditions, i.e., autonomous contributions that are not translations of the primary participants' utterances (Wadensjö, 1998). For instance, when patients ask for more information or for clarification, some interpreters tend to not relay the question but provide an answer to the question themselves (e.g., Amato, 2004; Cirillo, 2012). Furthermore, interpreters autonomously initiate questions or topics during the information exchange process (e.g., Amato, 2004; Bolden, 2000; Cirillo, 2012) and autonomously add empathic opportunities when relaying patients' utterances, which also prompt empathic responses, such as acknowledgements of the patients' feelings, from the HCPs (Theys et al., 2023). As such, these studies thus move away from the idea that interpreting is merely about producing accurate renditions and rather suggest that interpreters are active co-participants that engage in the information exchange process and in establishing a good doctor-patient relationship.

1.2. The potential impact of shifts in the interpretation process

Researchers who conceptualize interpreting as a process of strategic-decision making, and thus see interpreters as active co-participants, reflect on the impact of shifts on the final interpretation and suggest that it is not possible to classify omissions or other types of changes to the originals as systematically good or bad for the coordination of mutual understanding

and the coordination in the interaction more generally (e.g., Angelelli, 2004, 2019; Cirillo, 2012; Major & Napier, 2012; Wadensjö, 1998). Rather, whereas some shifts might lead to miscues, others might promote message equivalence and mutual understanding. For example, although omitting empathic information in the interpretation process might sometimes prevent the primary participants from establishing a good doctor-patient relationship (e.g., Gutierrez et al., 2019; Hsieh, 2016; Theys et al., 2023), some omissions might be strategic in nature. Interpreters can, for instance, opt to not relay empathic communication and thus restrict themselves to information they deem to be relevant for diagnostic purposes (e.g., Amato, 2004; Bolden, 2000; Cirillo, 2012; Davidson, 2000).

Moreover, interpreters can choose not to relay empathic information to avoid a potential misunderstanding. Theys et al. (2023), for instance, found that interpreters tend to relay HCPs' empathic responses to patient-initiated empathic opportunities, i.e., verbal expressions of emotion, challenge or process (Bylund & Makoul, 2002) as a close match. Doctors' empathic responses, which go from denial of the empathic opportunity to the doctor and patient sharing a feeling or experience (Bylund & Makoul, 2002), to interpreter-initiated empathic opportunities, however, are often omitted or reduced. Theys et al. (2023) suggest that when interpreters add a verbal expression of emotion or challenge to the interpretation of the patient's original and thus initiate an empathic opportunity, they might deliberately choose to omit or reduce the doctor's response to this empathic opportunity. This omission or reduction can be regarded as being strategic in nature, because patients might not relate to the emotions, challenges or progress introduced by the interpreter. Consequently, even though relaying the doctor's responses to these interpreter-introduced empathic opportunities might lead to more accurate renditions, they might cause misunderstandings and disrupt the coordination of mutual understanding. Thus, omitting and reducing empathic communication in the interpretation process can be seen as a strategy to optimize the HCPs' and patients' mutual understanding of empathic communication and promote a good patient-doctor relationship (Theys et al., 2023, p. 57).

1.3. owards a multimodal approach to the analysis of the information exchange process in healthcare interpreting

As sections 1.1 and 1.2 have shown, a large body of research highlights that (healthcare) interpreting entails strategic decision making based on a variety of contextual factors. However, when reviewing the literature, most of these studies have mainly focused on the analysis of verbal utterances. This is striking because research has shown that interlocutors use a variety of bodily resources when engaged in interaction and that they combine speech with other semiotic resources in the creation of larger composite utterances that prompt meaning construction (e.g., Clark, 1996; Enfield, 2009, 2013; Kendon, 2004; McNeill, 1992). These gestures serve a variety of functions, including referential, performative, modal and discursive functions (e.g., Müller 1998; Müller et al., 2013).

What is of particular importance here, is that gestures, especially representational gestures, can not only be co-expressive, but they can also add meaning that is not expressed in the verbal part of the utterance (e.g., Gerwing & Allison, 2009; Kendon, 2004; McNeill, 1992; Rowbotham et al., 2011). In a study on pain descriptions, for instance, Rowbotham et al. (2011) show that speakers frequently use gestures, and specifically representational gestures when talking about past pain experiences. Zooming in on the semantic speech-gesture interplay, it appears that a significant amount of information was expressed via gestures only or via speech-gesture composites. Information related to the location and size of the pain, for instance, was mainly captured in speakers' gestures only. Information related to the quality of the pain, however,

was expressed significantly more often via gesture-speech composites than in either the gestural, or the verbal mode only. Therefore, Rowbotham et al. (2011) not only suggest that gestures can add meaning onto the verbal parts of the utterance, but also that the creation of gesture-speech composites might be necessary to provide more accurate information.

Relating this back to the information exchange process and the coordination of mutual understanding in healthcare interpreting, these findings imply that omitting, modifying and/or adding gestures can impact the amount of information and thus the accuracy of the information that is transferred in the interpreting process in interpreter-mediated (medical) encounters. Indeed, also within the domain of Interpreting Studies there is an increased interest in the multimodal nature of face-to-face interaction. However, existing studies that investigate interpreters' and primary participants' bodily actions in dialogue interpreting mainly approach these in relation to multimodal interaction management and/or the creation of participation and engagement frameworks (see Davitti, 2019 for a recent overview). To the best of our knowledge, there is no study on how non-verbal shifts in interpreters' renditions can impact the amount and quality of the information exchanged between the primary participants in dialogue interpreting.

2. Positioning and aim of this paper

The current paper is part of a larger study that zooms in on multimodal shifts in the interpretation process, part of which has been presented at the IPrA 2023 conference. It appears that a substantial amount of the omitted gestures were of the representational type, which we—based on McNeill (1992)—defined as manual gestures and bodily enactments that refer to persons, objects, locations or events. These gestures might in some contexts add information to the verbal part of the composite utterance (cf. Kendon, 2004; McNeill, 1992). Consequently, gestural omissions, additions or modifications might lead to changes in meaning and can thus impact the information exchange process and the coordination of mutual understanding.

In the context of dialogue interpreting, gestures produced by one of the primary participants are often also visible to the other primary participant, i.e., to the addressee of the utterance. In the context of our data, this would imply that patients could perceive the HCPs' embodied behavior and map their meanings onto the verbal referents of the composite utterance when the interpreters translate the HCPs' utterances. As such, they could capture the full meaning of the HCP's composite utterances even when the interpreter does not repeat the HCP's gestures. However, as healthcare interpreters are often interpreting consecutively (cf. Pöchhacker, 2022), there can be a large temporal distance between the gestures in the originals and the interpretation of the verbal referents in the patient's mother tongue. This temporal gap makes it more difficult for the patient to semantically integrate the information provided via the spoken words and gestures and thus to capture the full meaning of the composite utterance (Özyürek, 2014).

The current paper addresses the potential impact of gestural shifts by providing a qualitative analysis of gestural omissions and additions in authentic interpreter-mediated interactions. We adopt a descriptive approach and do not consider additions and omissions as systematically good or bad, but we rather start from the idea that interpreting entails a process of strategic decision-making and thus from the idea that omissions or additions might even be recommendable in order to coordinate mutual understanding in dialogue interpreting (cf. Major & Napier, 2012; Wadensjö, 1998). In our analyses, we will thus consider how gestural omissions and additions can potentially affect the quality and amount of information being exchanged and the coordination of mutual understanding in healthcare interpreting.

3. Methodology

3.1. Dataset

As highlighted above, this paper takes findings from Beukeleers et al. (2023) as a starting point to further explore how gestural shifts can impact the amount and type of information that is being provided by healthcare providers. We selected three excerpts from the same authentic interpreter-mediated medical consultations. These consultations were recorded in an urban hospital in Flanders, the Dutch-speaking part of Belgium. The data were collected as part of the recently concluded project "Empathic Care for All" (Theys, 2021). In the consultations, a healthcare provider and foreign language-speaking patient communicated through a professional interpreter. The consultations were recorded at the departments of gynecology or endocrinology of the hospital. The interpreters were all trained and certified by an independent translation and interpreting agency funded by the Flemish government (*Agentschap voor Integratie en Inburgering*) and were hired by the hospital on a freelance basis. Before the consultation participants received informed consent in their native languages. The patients' informed consent forms were translated by professional translators. The duration of the consultations varied from 15 to 38 minutes. The study was approved by the hospital ethics committee (Belgian registration number: B322201835332).

The excerpts selected for this paper were taken from consultations with a Turkish-speaking patient and a Russian-speaking patient. None of the HCPs in this study were able to communicate in the patient's mother tongue. Patients reported that their language proficiency in Dutch varied from very limited to average. All patients and HCPs had already participated in an interpreter-mediated medical encounter before. The consultations were all first encounters between the patient and that particular HCP. However, in two consultations the interpreter and the patient had already met during a previous consultation with another HCP.

3.2. Transcription and translation

Professional translators—who were also native speakers of Russian and Turkish—transcribed the data and translated it into Dutch. Subsequently, translations were also revised by lecturers in the Linguistics Department of KU Leuven. For the purpose of the current study, the HCPs' utterances and the interpreters' renditions thereof were annotated in the ELAN annotation tool (Wittenburg et al., 2006).

3.3. Methods

In order to explore the impact of multimodal shifts, we identified all HCPs' utterances in which medical information, i.e., information related to the patient's illness and/or treatment (De Haes & Bensing, 2009), was conveyed to the patient as well as the interpreters' renditions of these utterances.

3.3.1. Identifying gestures

To be able to analyze multimodal shifts in our dataset, we first identified the HCPs' gestures. The beginning of a gesture was defined here as the onset of the preparation phase and the onset of the retraction phase was considered the end of a manual gesture (cf. Kita et al., 1998). For this paper, we aim to explore the impact of gestural omissions and additions in the interpretation process on the amount and type of information that the patient receives. Therefore, we focus on representational gestures, which we define as manual gestures and bodily enactments that refer to persons, objects, locations or events. They include iconic, deictic and specific types of metaphoric gestures (McNeill, 1992):

- Iconic gestures: imagistic gestures that depict formal characteristics of the person, object, location or event they refer to.
- Metaphoric gestures: imagistic gestures that depict abstract referents, such as knowledge, language or time.
- Deictic gestures: pointing gestures that indicate persons, objects, locations or events in the immediate environment or indicate non-present referents that are associated with a location in the gesture space.

When HCPs used a pen or the cursor on the computer to indicate a particular referent, these actions were coded as deictic gestures.

In case one gesture exhibited properties of two or more different categories, we aimed to identify the main function of the gesture within that particular context and annotated the gesture accordingly.

In this analysis, we also annotated segments during which the HCP used other artifacts to visually represent, i.e., to depict the medical information. These segments were annotated on the same tier and marked as:

- Drawing (e.g., drawing or showing a picture of an organ)
- Manipulating an object (e.g., folding a paper to depict a part of the treatment)

Finally, we created another tier to identify the interpreter's manual gestures and bodily enactments. They were annotated according to the same procedure as described above.

3.3.2. Identifying gestural shifts

We compared the HCPs' composite utterances with the interpreters' composite renditions in order to identify gestural shifts in the interpretation process. We thus compared HCPs' and interpreters' renditions both in terms of speech and in terms of embodied behavior in order to identify different types of gestural shifts and shifts in meaning in this study. We thereby adopted an inductive approach and established different types of shifts as they occurred in the data. In this paper, we selected excerpts that contained omissions and/or additions of representational gestures as these types of shifts were omnipresent in our data. We define these types of gestural shifts as follows:

- Omissions: HCP produced a manual gesture or enactment, but there is no equivalent gesture in the interpreter's rendition. Thus, the interpreter did not use the same or a similar gesture with the same function, nor did he/she verbalize the information captured in the gesture.
- Additions: manual gestures or enactments that were introduced autonomously by the interpreter for which there was no equivalent present in the HCP's original.

Note that we also considered the surrounding speech and other bodily actions when identifying the different types of shifts. In theory, interpreters could also verbalize information that was communicated via gestures only in order to relay the information. However, this did not occur in our dataset.

3.3.3. Analyzing the impact of a gestural shift

Subsequently, the three examples of gestural omissions or additions presented in this paper are analyzed for how the shift potentially impacted the information provided to the patient. In order to investigate this, we identified gesture-speech composites in the HCPs' utterances and analyzed which piece of information was described, indicated, and/or depicted in each modality (cf. Clark, 1996; Enfield, 2009). We define these different methods as follows (e.g., Clark, 1996; Enfield, 2009):

- Describing: communicating a referent by telling, i.e., representing it categorically (e.g., referring to the location of the surgery with the linguist category of 'blatter').
- Indicating: anchoring a referent to the real world by locating it in space and time (e.g., indicating the location of the surgery with indexicals like "it", "there", pointing gestures or eye gaze).
- Depicting: communicating a referent by showing how it looks, sounds, or feels like (e.g., showing the location of the surgery by using an iconic gesture that depicts the organ or by drawing the organ on a piece of paper) .

Note here that these are methods of communication and that they cannot be easily distinguished from one another in actual language use. Speakers combine these methods in the creation of composite semiotic signs (Clark, 1996; Enfield, 2009). Thus, one word, one gesture or one utterance often reflects different methods simultaneously and/or sequentially.

After analysing the HCPs' utterances, we identified the equivalents in the interpreter's renditions. We investigated whether, and if so, which piece of information was modified or omitted in the interpretation process as a result of the gestural omission. In a similar vein, we identified speech-gesture composites in the interpreters' renditions that were annotated as additions and analyzed them in the same manner as the composite utterances with gestural omissions, i.e., we compared the HCP's original and the interpreter's renditions thereof in terms of describing, depicting and/or indicating information.

4. Analyses

In this section, we present the three examples of gestural shifts and elaborate on their impact on the information exchange process and the coordination of mutual understanding. The examples in 4.1 and 4.2, respectively, show how interpreters omit or add iconic gestures. The final example illustrates how omitting deictic gestures can lead to repair initiation, i.e., an interlocutor signaling difficulties in understanding (part of) the previous turn-at-talk (Schegloff, 2000) and how chunking can aid in relaying visual information to the patient (4.3).

4.1. How gestural omissions can impact the amount and type of information exchanged

The first excerpt is taken from a medical encounter with a Russian-speaking patient, a Dutch-speaking HCP and a professional interpreter. The patient had already visited the department of Endocrinology, and surgery to remove the thyroid was scheduled. However, the patient had decided to cancel the surgery. She now returns to the hospital and at the start of this consultation, it appears that the patient had not fully understood the result of the puncture taken during the previous consultation. She seems to be in doubt about whether the lump in her thyroid is malignant or not. Therefore, the HCP is explaining how a puncture works and why surgery is recommended. In doing so, the HCP uses a variety of representational, often iconic, gestures that depict a puncture and the analysis of the cells taken during this procedure. These gestures are, however, omitted by the interpreter. In this section, we zoom in on the possible impact of the shift on the amount of information that the patient receives.

Excerpt 1[1]

1. **HCP**	*Dus (0.8) we hebben (0.7) als wij een (0.9) knobbel zien uhm die er verdacht*
2. **(00:06:36)**	***So (0.8) we have (0.7) if we see a (0.9) lump that looks uhm suspicious,***
3.	**#Fig. 1----------**
4.	*uitziet, gaan we daar in prikken. (0.5) En dan (0.9) kunnen we nooit 100%*
5.	***we are going to prick it. (0.5) And then (0.9) we can never be a 100%***
6.	**#Fig. 2----------**

[1] We have included transcription conventions in the appendix of this paper.

7. *zeker zeggen op basis van die punctie alleen (.) of het nu echt kanker is of niet.*
8. ***certain based on the puncture only (.) whether it is cancer or not.***
9. *Ma we kunnen daar wel een graad van verdachtheid uit afleiden en bij haar*
10. ***But we can use it to determine a degree of suspicion and in her case (.)***
11.
12. *was er een hoge verdachtheid (.) uhm dat het mogelijks kwaadaardig kan zijn*
13. ***there was a high degree of suspicion (.) uhm that it can possibly be malignant.***

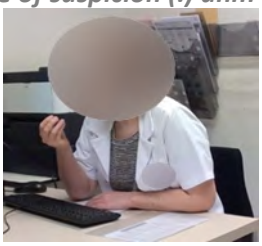


Figure 1. HCP producing an iconic gesture to depict "puncture"

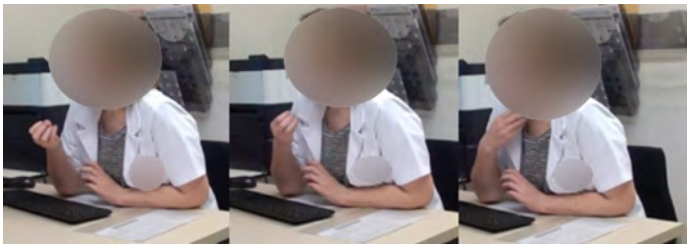


Figure 2. Reintroduction of the iconic gesture for "puncture"

54. **I** Когда мы видим узело (0.5) узелок ээ нам кажется подозрительным (.)
55. **(00:07:41)** ***When we see a lum (0.5) a lump that u::h looks suspicious to us (.)***
56.
57. мы берем пункцию и никогда пункция не может дать сто процентов
58. ***we do a puncture and never can a puncture give a 100%***
59. **#Fig. 3-----** **#Fig. 4------**
60. гарантии злокачественно это или не злокачественно (.) мы берем
61. ***certainty that it is malignant or not malignant (.) we take***
62.
63. только несколько (.) клеток и на основе этих клеток (.) этого результата (.)
64. ***only a few (.) cells and on the basis of these cells (.) of this result (.)***
65.
66. мы не можем рисковать чтобы говорить это сто процентов так или
67. ***we cannot risk to say that it is 100% like this***
68.
69. это сто процентов не так но (0.5)
70. ***or that it is 100% like that but (0.5)***
71.
72. мы можем степень подозрительности все-таки более менее определить
73. ***we can still determine more or less the degree of suspicion***

74.
75. у Вас она была достаточно подозрительна
76. ***and with you it was sufficiently suspicious***

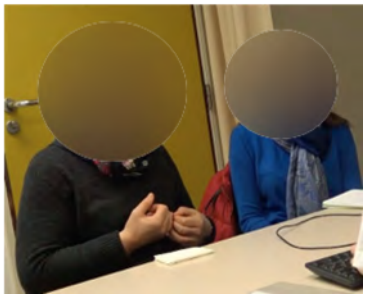


Figure 3. Interpreter (on the right) during her interpretation of the medical procedure "puncture"

Figure 4. Interpreter (on the right) during her interpretation of "puncture"

In lines 1-8, the HCP explains that they tend to do a puncture if they find a lump that looks suspicious. The HCP starts with *"So we have"* (line 1), pauses briefly and subsequently restarts with *"if we see a lump that looks uhm suspicious, we are going to prick it"* (lines 1-5). Zooming in on her bodily movements, she initiates an iconic gesture as a strategy to depict the puncture in line 1 (Fig. 1), which is interrupted as she produces a restart of her utterance but is reintroduced during *"we are going to prick it"* (line 5, Fig. 2). The HCP's right hand enacts the holding of the needle and thus depicts what the procedure looks like (Müller, 2014). Moreover, as she moves her hand to her neck, she also uses her own body to depict where the patient has had a puncture, i.e., she indicates that a puncture was taken from the thyroid (Fig. 2). Whereas the verbal part of the utterance thus accurately describes the action performed (pricking the lump), the gesture adds a depiction of what the action looks like and an indication of the exact location (cf. Clark, 1996, 2016; Enfield, 2009).

When comparing the HCP's original utterance with the interpreter's rendition, we see shifts on both the verbal and non-verbal level. First, the interpreter modifies *"we are going to prick it"* with "puncture", rather than using a similar simplified explanation (line 57). Moreover, the interpreter leaves her hands on her lap during her interpretation and does not produce any (clearly visible) gestures here (Fig. 3 and 4). This implies that the visualization and the reference to the thyroid in the HCP's original are not rendered by the interpreter and thus that the HCP's speech-gesture composite entails more information and also uses more modes of representation, i.e., she also visualizes the information.

Both the HCP's bodily actions, i.e., the visualization, and the verbal counterpart of the utterance contain cognitive aid strategies (cf. Menichetti et al., 2021). The HCP first provides a simplified explanation of the medical procedure, i.e., "pricking the lump", before introducing the medical term "puncture" (line 8). Next, while she pronounces *"based on the puncture"*, the HCP repeats the iconic gesture she previously used to explain what a puncture is and reduplicates its movement until the end of "puncture" (line 8, Fig. 3). Thus, even when reintroducing the referent, she does not only use the more technical term "puncture", but also repeats the depiction thereof.

In sum, this excerpt shows that omitting an iconic gesture in the interpretation process can lead to changes in the amount and type of information that is being provided to the patient. Whereas the HCP combines the verbal description of the puncture with a vivid depiction of this particular medical procedure and with an indication of the location, the interpreter only renders the verbal description. Moreover, as the interpreter has omitted both the verbal simplification for "puncture" and the visualization of the medical procedure, which can also be regarded as a cognitive aid strategy (cf. Menichetti et al., 2021), it might be more demanding for the patient to capture the full meaning of "puncture" based on the interpreter's rendition than based on the HCP's explanation. We will return to this in the discussion (cf. section 5).

4.2. How adding iconic gestures can result in more concrete renditions

In the second example, we zoom in on an excerpt in which the interpreter adds an iconic gesture during her interpretation and, in doing so, makes the original utterance more concrete. This example is taken from a consultation at the department of Endocrinology. The HCP, a Turkish-speaking patient and a professional interpreter are engaged in an encounter about surgery to remove an adenoma in the patient's pituitary. At this point, the patient is, however, reluctant, as he had an operation in Turkey already. As they did not manage to remove everything, and he still had many health issues afterwards, he is not sure whether additional surgery would solve his issues.

Excerpt 2

1. **HCP** In iedere operatie (.) er zijn altijd risico's verbonden (.) ik weet niet (.) wat er
2. ***Each surgery (.) comes always with certain risks (.) I don't know (.) what***
3. gebeurd is in Turkije want daar bent u geopereerd (.) Ik zag dat u ook (.) wat euh
4. ***happened in Turkey because you had your surgery there (.) I also saw that (.) u::h***
5. euh lekkage heeft gehad (.) euh waardoor dat u waarschijnlijk wat afgezien heeft
6. ***u::h you also had a leak (.) u::h and because of that u have probably been suffering***
7. (0.9) maar (0.5) euhm (1.5) het is wel zo (0.6) dat als (0.7) euh (1.2) als wij
8. ***(0.9) but (0.5) uhm (1.5) it is a fact (0.6) that if (0.7) euh (1.2) als wij***
9. voorstellen (.) om te gaan kijken om het te opereren (.) euh het is voor
10. ***suggest (.) to take a look, to do the surgery (.) u::h it is the aim to do***
11. een vervolledigheid van de resectie (.) en we moeten dit ook (.) voorleggen
12. ***a full resection (.) and we also have to (.) present this***
13. aan de neurochirurgen (.) om te zien (.) als het (.) wel toch (.) een mogelijkheid
14. ***to the neurosurgeons (.) to see (.) whether it (.) is (.) a possibility***
15. (0.6)
16. Nu (1.9) we kunnen niet (0.8) we weten niet wat er gebeurd is in Turkije dus
17. ***Now (1.9) We cannot (0.8) we don't know what happened in Turkey so***
18. ik kan niet zeggen ja het was niet mogelijk om die volledig weg te doen ook niet
19. ***I can't say that yes, it was not possible to fully remove it or not***
20. Da weten we nie e
21. ***We don't know that huh***
22. ***(1.7)***
23.
24. **I** şimdi (.) her ameliyat riskli (.) hı=
25. ***now (.) each surgery is risky (.) hu=***
26. **P** =e tabii ki=
27. ***=uhu of course***
28. =ama biz burada diyosak hani ameliyat ol diye he tabii bunu ameliyat eden
29. ***=but if we say here, you know, do the surgery hu (.) then of course the***
30. doktorun da görmesi gerekiyor önceden (.) ama buradaki amaç hepsini almak
31. ***doctor that operates has to see this beforehand (.) but the aim is to remove it all***

32. (.) hi (.) şimdi Türkiye' de ne Iduğunu bilmiyorum çünkü hani sonradan
33. ***(.) hu (.) I don't know what happened in Turkey because you know after the***
34. ameliyattan sonra akıntı o:lmuş falan hani
35. ***surgery you had some discharge***
36. **#Fig. 5**------
37. baya rahatsızlık olmuşsun ama neler olduğu bilmediğim için Türkiye' de (.)
38. ***which was really disturbing but because I don't know what happened in Turkey***
39. [bir şey diyemiyorum
40. ***I can't say anything about that surgery***

Figure 5. Interpreter adding an iconic gesture that depicts "discharge"

While explaining that each surgery comes with risks, the HCP refers to the patient's surgery in Turkey. In line 5, the HCP mentions that it is indicated in the patient's medical record that he had a leak from which he was probably suffering. In her utterance, however, it is unclear what type of leak the HCP exactly refers to. This contrasts with the interpreter's rendition. In line 35, we see that she also uses a broad term to refer to the patient's health condition with "*you had some discharge".* However, in contrast to the HCP, the interpreter also produces an iconic gesture that depicts the discharge (Fig. 5). She uses an open hand that starts at her mouth and moves away from her body. In this way, the interpreter depicts the movement of the discharge and indicates that it came from the patient's mouth by pointing at her own mouth at the start of gesture (cf. Clark, 1996; Enfield, 2009). This might indicate that the patient had to vomit often after the surgery or that he would throw up blood. In that way, the interpreter does provide more specific information to the patient than the HCP, i.e., she narrows down the options of types of discharge. At this point in the encounter, the patient had not mentioned this symptom. Later in the encounter the patient also only mentions that he has been suffering from a runny nose, but he does not refer to vomiting or any other type of discharge that could be related back to this iconic gesture. Thus, by adding an iconic gesture that specifies a particular type of discharge, the interpreter renders a more specific composite meaning that is potentially wrong.

If the patient indeed often had to throw up after the surgery and if the HCP is indeed referring to that particular complaint at this moment in the interaction, one could argue that adding the iconic gesture is an efficient interpretation strategy that results in a more concrete rendition and thus facilitates the information exchange process and the coordination of common ground. However, as highlighted above, there is no reference to "throwing up" as a complaint in the entire consultation. The interpreter in this consultation was present in previous consultations with this patient and another HCP at the department of Endocrinology. As such, she has prior knowledge of the patient's health condition and she might thus have learned about this complaint in a previous consultation and use this information in her interpretation at this moment in the consultation. In that case, it could be that the addition of the gesture does not necessarily lead to an interpretation error. However, even then, there is no interactional evidence that the HCP is referring to this particular complaint or another in the excerpt above. It thus remains uncertain whether the rendered composite utterance is correct.

In sum, the excerpt above illustrates how interpreters can render more concrete meanings compared to the meaning of the primary participant's original by adding an iconic gesture that depicts a medical symptom. As gestural addition can contribute to more specific information and to visualization, one could argue that adding iconic gestures can facilitate the negotiation of meaning and that they promote the coordination of mutual understanding within a particular interaction environment. However, as illustrated above, making information more concrete and/or visualizing information comes with certain risks as it can potentially lead to incorrect composite utterances, and thus to errors in the information exchange process, as well. We return to this in the discussion.

4.3. Repair initiation after omitting deictic gestures

The final example is taken from the same consultation at the department of Endocrinology as excerpt 1 (cf. 4.1). At this moment, the HCP in this consultation updates the professor and brings her in to talk to the patient. The professor encourages the patient to raise her questions and concerns, but first wants to recapitulate the results from the ultrasound diagram and the puncture that were taken during the previous consultation. She does this by showing the ultrasound diagram and explaining what they have found. In this excerpt, we focus on the deictic gestures and the process of mapping verbal meanings onto the visual referents on the ultrasound diagram only.

Excerpt 3

1. **HCP** *Dus (.) euhm (2.5) dit is hier de luchtpijp (2.6) de luchtpijp e (0.9) en hier is*
2. ***So (.) uhm (2.5) this is here the trachea (2.6) the trachea huh (0.9) and here is***
3. **#Fig. 6-------**
4. *eigenlijk de rechterkant van de schildklier (1.0) en het witte gedeelte is eigenlijk*
5. ***actually the right side of the thyroid (1.0) and the white part is actually***
6. **#Fig. 7---------- #Fig. 8---------**
7. *normaal (0.9) maar heel (0.8) die inliggende donkere zone (.) is eigenlijk het gezwel*
8. ***normal (0.9) but this entire (0.8) internal dark area (.) is actually the tumor***
9. **#Fig. 9----------**
10. (1.6)
11. *Misschien kan je dat al [effe]kes vertalen?*
12. ***Maybe you can already translate this for a bit?***
13. **I** *[ja]*
14. ***[yes]***
15. То есть вот здесь у Вас проходит трубка
16. ***So here is the tube through***
17. через которую мы дышим. Здесь с правой стороны (.) светлое белое место (.)
18. ***which we breathe (.) Here on the right side (.) is a bright white spot (.)***
19. Это нормальная часть Вашей щитовидки
20. ***That is the normal part of your thyroid***
21. **P** Это вот это да? Нормальная..
22. ***That is this then, right? The normal...***
23. **#Fig.10----------**
24. **HCP** *Dit is normaal*
25. ***This is normal***
26. **#Fig. 11----------**
27. (0.5)
28. **I** да
29. ***Yes***
30. **#Fig. 12----**
31. **HCP** *Dit is luchtpijp*

32. ***This is trachea***
33. **#Fig. 13----------**
34. (0.6)

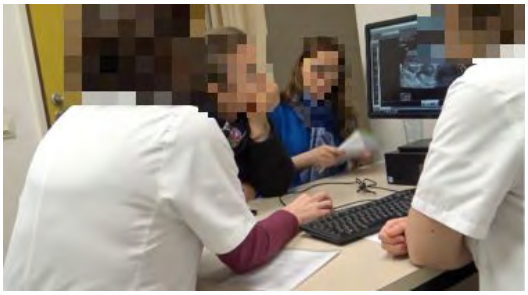


Figure 6. Professor indicating the trachea

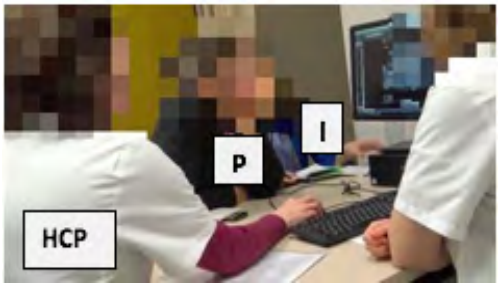


Figure 7. Professor indicating the right side of the thyroid

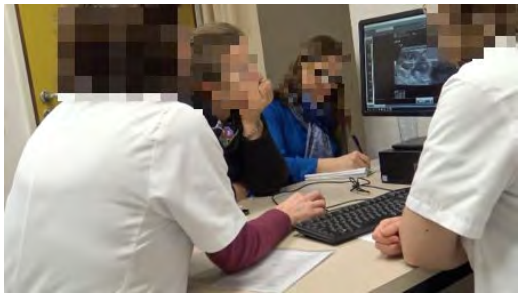


Figure 8. Professor indicating the white part

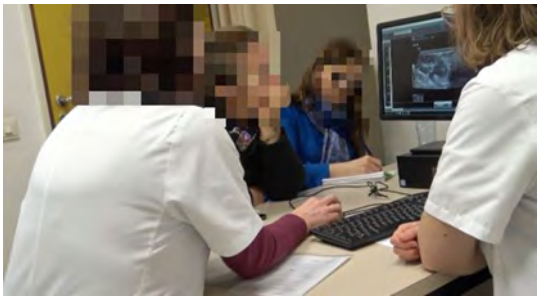


Figure 9. Professor indicating the tumor

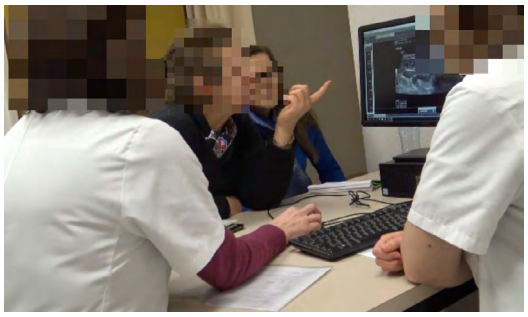


Figure 10. Patient initiating repair with a deictic gesture

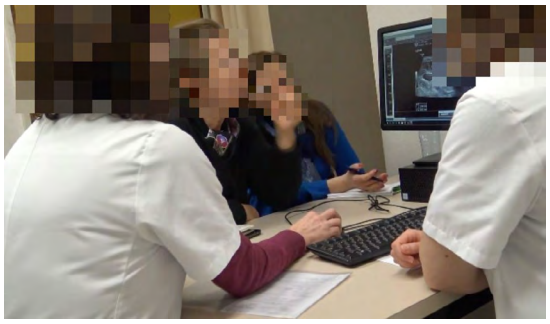


Figure 11. Professor again indicating the normal part of the thyroid

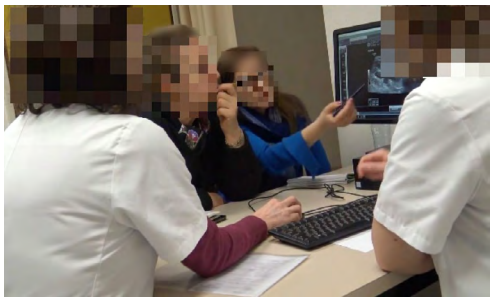


Figure 12. Interpreter relaying the professor's repair turn and indicating the normal part of the thyroid

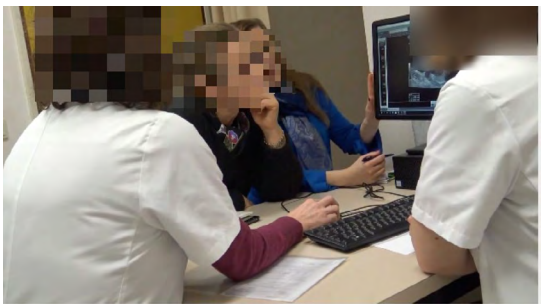


Figure 13. Professor indicating the trachea

35. **P** *«Luchtpijp» что ma[кое?*
36. ***«Trachea» what is [that?***
37. **I** [Это то через которое мы дышим с Вами только
38. [***That is what you and I need to breathe through, [but***
39.
40. **HCP** [*Dit zijn*
41. ***[These are***
42. **#Fig. 14----**

43. *normale bloedvaten*
44. ***normal blood vessels***
45. ---
46. **I** Это нормальные кровяные сосуды
47. ***These are normal blood vessels***
48. **HCP** *En dit is eigenlijk de rechterkant van de schildklier*
49. ***And this is actually the right side of the thyroid***
50. **#Fig.15----------**
51. **I** Это правая часть Вашей щитовидки
52. ***That is the right side of your thyroid***
53. **HCP** *Waarbij dat dat lichtgrijze nog normaal is maar het donkergrijze is eigenlijk het*
54. ***And the light grey is still normal (.) but the dark grey is actually the***
55. **Fig. 16----------** **#Fig. 17---------**
56. *gezwel waarin we geprikt hebben*
57. ***tumor which we pricked***
58. **I** Светло-серая часть это еще нормальная,
59. ***the light grey part that is still normal (.)***
60. темная часть, это та часть в которой мы брали пункцию
61. ***the dark part, that is the part in which we took a puncture***
62. (1.5)
63. **P** Это значит вот это
64. ***That is then that part?***
65. **#Fig.18----------**
66. **I** *Dat is hier dus?*
67. ***So, that is this here?***
68. **----------**

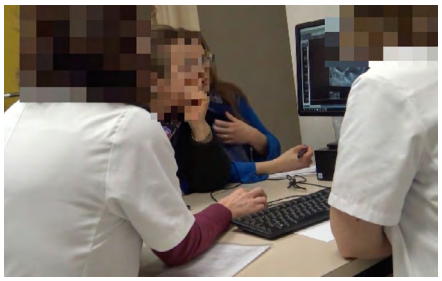


Figure 14. Professor indicating the normal blood vessels

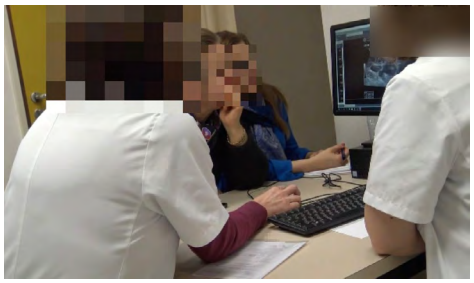


Figure 15. Professor indicating the right side of the thyroid

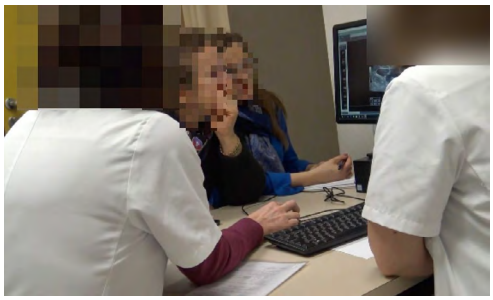


Figure 16. Professor indicating light grey area

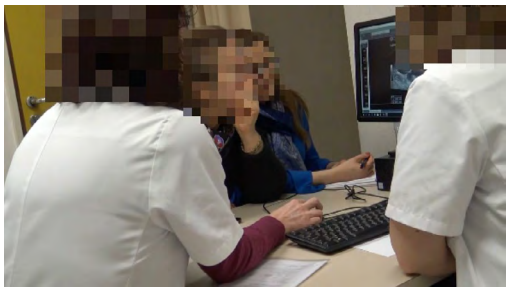


Figure 17. Professor indicating dark grey area

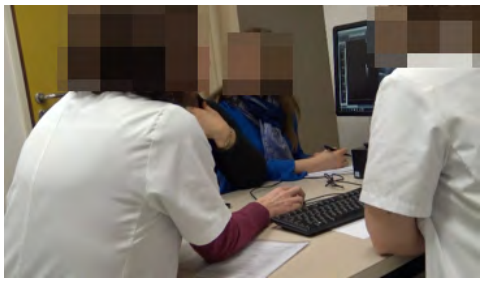


Figure 18. Patient indicating thyroid on her own body

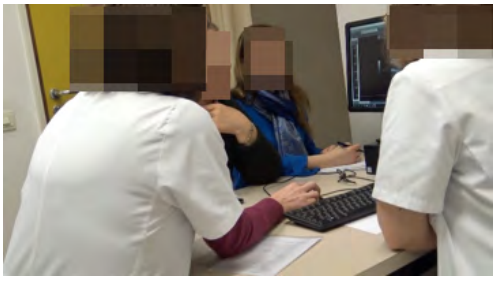


Figure 19. Professor indicating the needle

69. **HCP** *Dus ja en dit is eigenlijk de naald hier van de vorige punctie*
70. ***So yes and this is actually the needle here from the previous puncture***
71. **#Fig.19----------**
72. **I** Это здесь иголка из прошедшей, из прошлой пунк[ции]
73. ***This is here the needle from the past, from the previous puncture***
74. **----------**
75. **HCP** [ja?]
76. ***[yes?]***
77. *En dus ja d-die e daar dat staal dat toen is opgestuurd naar het labo (.)*
78. ***And so yes, t-that uh that sample that was then sent to the lab (.)***
79. *ik denk dat we da (.) vorige keer ook ook met u besproken hebben (.) das*

80. ***I think that we (.) the last time also have discussed this (.) all in***
81. *in het Nederlands allemaal maar ik wil het toch nog eens tonen aan u dat u echt*
82. ***Dutch (.) but I want to show you that one more time so you are really***
83. *overtuigd bent (.) hier staat besluit e (0.7) papillair schildklier carcinoom (.) [dat*
84. ***convinced (.) here is the conclusion hu (0.7) papillary thyroid carcinoma (.) [that***
85. betekent] dit is carcinoom
86. ***means] this is carcinoma***
87.
88. **P** [carcinoom
89. ***[carcinoma***
90. *ja]*
91. ***yes]***

In lines 1-12 the HCP explains what they can see on the ultrasound. During her explanation, she first introduces the trachea (line 1). While describing the referent verbally, she traces its shape and indicates its location on the ultrasound with the cursor of her mouse (Fig. 6). Subsequently, the professor traces the right side of the thyroid on the echography, explaining that the white part is normal and that the entire dark area is the tumour (Fig. 7-9). The cursor movements as such do not only indicate the referents on the ultrasound diagram but also depict their size and shape by tracing them. The professor then pauses, however, as the interpreter does not initiate a turn just yet, she self-selects again and explicitly asks the interpreter to relay this information already (line 12).

The interpreter meets the request and relays the professor's explanation (line 15-19). Most relevant for our analysis is the fact that the interpreter does not use any (deictic) gestures to indicate the referents on the echography. The only indexical elements in her rendition are the verbal elements (вот (*here*), Здесь (*here*), Это (*this*)). However, a physical connection between the verbal and the visual referents is lacking during the rendition, which might make it more difficult for the patient to establish the right meaning mappings. Note that this could be related to the fact that the interpreter was taking notes during the professor's explanation and thus was looking in the direction of her booklet, rather than in the direction of the ultrasound diagram (cf. Fig. 6-9).

The idea that the absence of physical points towards the visual referents might lead to difficulties in the grounding process, i.e., in the coordination of mutual understanding, becomes apparent in line 22. The patient initiates repair immediately after the interpreter's renditions, i.e., she signals that she could not understand part of the information by asking *"That is this then, right? The normal …"* (Schegloff, 2000). While raising the question, the patient also simultaneously traces part of her thyroid on the ultrasound diagram, i.e., she thus indicates and depicts the referent she couldn't understand on the computer screen (Fig. 10). As such, the repair initiation supports the idea that the patient had difficulties with mapping the verbal onto the visual referents. In what follows, the interpreter—as the speaker of the turn with the trouble source—does not relay or immediately reply to the question. Rather, the professor understands that the patient is having trouble with identifying the referents on the ultrasound diagram and immediately provides a repair turn herself. She repeats *"this is normal"* (Fig. 11). Note that this time, all participants are oriented towards the computer screen and the interpreter is not taking notes. Rather, the professor pauses briefly and provides the interpreter the time to relay this brief segment. After a brief pause (0.5), the interpreter confirms to the patient that this area is normal, by replying *"yes"* to the patient's question and by indicating the location with a deictic gesture (line 32, Fig. 12).

In what follows, the professor repeats the different referents visible on the ultrasound diagram introduced earlier in a similar vein, i.e., by verbally describing them and by indicating them

on the screen with her cursor. However, this time she chunks the information into smaller segments, i.e., she pauses after each referent to provide the interpreter with the opportunity to relay the information immediately. The interpreter relays each referent again only by translating the verbal utterance. However, as the information is chunked into brief segments and the cursor still indicates the referent, this does not seem to cause any misunderstanding.

Only in line 63, the patient again initiates repair. As she asks *"That is this then, it's that part?"*, she uses a deictic gesture to indicate the location of her thyroid on her own body (Fig. 18). Thus, the repair initiation does not indicate difficulties with the mapping on the echography diagram, i.e., does not indicate the absence of a pointing gesture or another physical deictic element in the interpreter's turn as the trouble source, but rather the absence of the mapping of the referents onto her own body. The interpreter relays the patient's repair initiation and, subsequently, the professor provides a repair by confirming with "yes" and by indicating the needle of the puncture on the ultrasound diagram (line 69, Fig. 19).

The interpreter immediately relays the repair turn (line 72). She does not produce a gesture to indicate the needle on the echography, as the professor does. However, the cursor is still visible and the interpretation again considers a brief chunk of information with one referent. Consequently, the gestural omission does not seem to be problematic. This is supported by the fact that the patient releases her deictic gesture on her thyroid when she hears the interpretation, indicating that the repair has succeeded. Moreover, towards the end of the interpreter's rendition, the professor explicitly checks for addressee comprehension by looking in the direction of the patient and asking "*yes?"*. The patient confirms by producing a non-verbal acknowledgment token, i.e., by nodding multiple times (Gardner, 2001). As such, the participants orient towards sequence closure (Schegloff, 2007). In the next line, the professor then initiates a new course of action by showing the conclusion of the analysis and explaining that it is for sure malignant.

In sum, this excerpt shows that omitting deictic gestures in the interpretation process can lead to difficulties in the interactional process of establishing common ground (Clark & Brennan, 1990). However, in the repair organization, we see that chunking seems to be an efficient cognitive aid strategy that allows patients to reconstruct meaning and interpreters to relay information without having to point to the ultrasound diagram themselves. We will return to this in the discussion.

5. Discussion

5.1. On the semiotic complexity of the information exchange process

We presented three excerpts in which HCPs and/or interpreters used iconic and/or deictic gestures that aided in the visualization of the medical information. The excerpts thus illustrate how both HCPs and interpreters do not only describe a medical procedure or symptom but also depict and/or indicate some aspects of that meaning. By visualizing the information, they often integrate information into their composite utterances that is not expressed verbally at all. In our examples, this mainly involved a deictic and/or an iconic feature, i.e., the indication of a location and/or the depiction of a particular movement. In a similar vein, the deictic gestures in the HCPs' utterances can aid in the visualization of the medical information. First, they are physical points that provide a cue on how to map the verbal referents onto their visual counterparts on the ultrasound diagram (cf. Clark, 1996; Enfield, 2009). Moreover, some of the deictic gestures also traced the size and shape of their referent and can thus be regarded as also adding a depictive element to the composite utterance. These examples thus support the idea that speech-gesture composites can be more concrete than information that is provided through speech only (cf. Clark, 1996; Enfield, 2009; Gerwing & Allison, 2009; Rowbotham et

al., 2011) and point towards the semiotic complexity of the information exchange process in healthcare settings and in healthcare interpreting. Moreover, as previous studies mainly described the use of pictures or images as visual cognitive aid strategies in medical settings (cf. Menichetti et al., 2021), the current study adds the use of representational gestures in this regard.

5.2. Gestural shifts and their potential impact on the information exchange process

When comparing the originals and the renditions, it appears that interpreters omitted iconic and/or deictic gestures, or added iconic gestures. The gestures depicted a medical treatment, a symptom, the size and shape of a referent and/or indicated locations (i.e., the location of organs on the body or on an echography). The analyses in this paper do not only illustrate that gestural shifts occur but also shed light on the potential impact of the shifts on the type and quality of information that is being exchanged and thus on the coordination of mutual understanding in more general. First, we have seen that gestural omissions and/or additions can lead to less/more concrete renditions. On the one hand, as interpreters often omit the representational iconic and deictic gestures, the visual information provided through them (e.g., size and shape, locations, enactments) is often not relayed. The second excerpt, on the other hand, shows that interpreters can also make information more concrete by adding iconic gestures. Thus, the analysis of excerpt 2 shows that—in dialogue interpreting—it is not sufficient to only/mainly include verbal analyses of interpretations in order to determine the degree to which an interpreter's rendition can be considered accurate (cf. Aranguri et al., 2006; Flores et al., 2003; Hsieh, 2016) or to describe (shifts in) the information exchange process (cf. Angelleli, 2004, 2019; Wadensjö, 1998). Rather, when investigating message equivalence and the coordination of mutual understanding in interpreter-mediated discourse, including visible bodily action in the analyses can yield different insights (cf. Angelleli, 2004, 2019; Theys, 2021; Theys et al., 2023; Wadensjö, 1998).

Moreover, the excerpts presented in this paper support the idea that omissions and/or additions of gestures cannot be systematically categorized as being either errors in the process or good interpretation strategies (cf. Major & Napier, 2012; Wadensjö, 1998). This is illustrated in the second excerpt, where the addition of an iconic gesture makes a particular symptom more concrete, i.e., it narrows down the types of discharge the patient can suffer from. On the one hand, the gestural addition can be an effective strategy for visualizing and/or providing more concrete information and thus as an effective strategy that promotes the coordination of mutual understanding. However, in this example there is no interactional evidence that "discharge" refers to the meaning of "vomiting" or "throwing up". Even when the interpreter relies on prior knowledge and his/her common ground with the patient, it remains unclear whether "vomiting" is the exact symptom that the HCP is referring to at this moment in the interaction. Therefore, the addition of the iconic gesture can potentially lead to an interpretation error and thus to difficulties in the information exchange process and in the coordination of mutual understanding.

Finally, the analysis of the third excerpt illustrates that the omission of deictic gestures can lead to repair initiations when the patient is provided with large chunks of information. However, when chunking the information and reducing the amount of information, it does not appear to be problematic to omit the deictic gestures. Therefore, the analyses in this paper support the idea that omitting, and/or adding (gestural) information in the interpreting process does not necessarily lead to interpreting errors and/or communicative troubles (cf. Angelleli, 2004, 2019; Cirillo, 2012; Napier, 2004; Major & Napier, 2012; Theys et al., 2023; Wadensjö, 1998). Rather, interpreting is an act of strategic decision-making, which implies that interpreters can

omit, modify and/or add information to the originals based on the interactional context and their understanding of the (medical) information in order to coordinate mutual understanding. This paper, however, adds the dimension of visible bodily action to this discussion.

5.3. Visualization and gesturing in light of "cognitive aid strategies"

As visualization is considered a cognitive aid strategy that HCPs use to simplify complex medical information (cf. Menichetti et al., 2021), the excerpts presented in this paper also suggest that—depending on whether the interpreter omitted and/or added gestures—interpreters' renditions can be either more or less difficult to comprehend compared to the HCPs' utterances. When considering gestural omissions, we acknowledge that patients in our examples are involved in face-to-face interaction and thus often have full visual access to the HCP's visible bodily action. Thus, they could perceive the HCPs' gestures and subsequently map them onto the interpretation, i.e., the verbal utterance of the interpreters to capture the full meaning of the provided information. However, as highlighted in section 2 of this paper, the temporal gap between the composite utterance of the HCPs and the interpretation thereof might make it more difficult for the patient to semantically integrate the information provided via the spoken words and the gestures (cf. Özyürek, 2014). In other words, it might make it more difficult to capture the full meaning of the composite utterances. This becomes particularly apparent in the third example of this paper, where we have shown that the omissions of deictic gestures that physically point towards visual referents can in some cases, especially when there is a large temporal gap between the original and the rendition, lead to difficulties in the coordination of common ground and, consequently, can lead to repair initiations. Furthermore, research has shown that gestures, and in particular speech-related iconic gestures, facilitate the automatic semantic integration of gesture and speech (Chui et al., 2018) and that addressees are significantly better at recalling and recounting information accurately when iconic gestures are available (Beattie & Shovelton, 2001). Relating this to an interpreter-mediated context, it thus appears that omitting or modifying gestures can make it more demanding for patients to process the information. Furthermore, the use of iconic, metaphoric, and/or deictic gestures can also aid the interlocutors with the semantic processing and the coordination of common ground (Chui et al., 2018). In that regard, omitting, modifying and/or adding gestures or visual input does not only relate to the notion of 'accuracy' and the quality of the information exchanged, but also to the use of cognitive aid strategies.

On the one hand, the excerpts analyzed in this paper can help interpreters to recognize visual cognitive aid strategies used by HCPs. On the other hand, the excerpts can also inspire interpreters to initiate visual communication strategies autonomously in order to facilitate the coordination of mutual understanding. As we have seen in excerpt 2, interpreters can visualize medical information by adding iconic gestures and, as visualization can be regarded as a cognitive aid strategy (cf. Menichetti et al., 2021), one might argue that this can be an efficient strategy for interpreters to promote a better understanding of the medical information. However, as discussed above, the additions can potentially lead to errors and healthcare interpreters are not medical experts themselves. Therefore, caution is always warranted (see also Major & Napier, 2012 on visualization as an effective interpreting strategy in Australian Sign Language).

5.4. Coping with the semiotic complexity of healthcare communication

In this paper, we have explored the semiotic complexity of healthcare interpreting and the impact of gestural shifts on the coordination of common ground. One factor that is worth considering in this discussion is the fact that interpreters in dialogue interpreting are often relaying consecutively and thus are also often involved in notetaking (cf. Pöchhacker, 2022). In the examples we discussed here, interpreters frequently engaged in notetaking and mainly

gazed at their booklet. This implies that they might not have had full visual access to the primary participants' visible bodily action, i.e., they might not have seen the gestures or have only registered them in their periphery view. Consequently, it might not be straightforward to integrate visual information in their performances. Therefore, it is not only important to raise awareness about the use of gestures (in healthcare settings) amongst interpreters but also to inform HCPs about the complexity of integrating visual information in the interpreting process. In that way, they can collaborate and seek more effective and efficient communication and interpreting strategies in order to ensure that the patients have full access to the complex composite utterances and thus that they receive the most optimal interpretation.

In the third example, the repair organization indicates that chunking can, for instance, be an effective strategy to cope with the semiotic complexity of healthcare communication. On the one hand, it reduces the cognitive load for interpreters (Huang et al., 2023) and, as the information is provided in brief chunks, they do not have to take notes. Thus, it allows them to look at the primary participants' bodily actions and integrate such visual information in their interpretation. On the other hand, it also reduces the cognitive load for patients (Menichetti et al., 2021), as chunking reduces not only the amount of information per chunk, but also the temporal gap between the HCPs' visible bodily action and the interpretation in the patients' mother tongue. This might also aid patients in processing speech and gesture automatically (Özyürek, 2014). In other words, it might allow them to still map a gesture onto its verbal meanings, even when the interpreter does not mirror the HCP's gesture, just like in our third example.

5.5. Limitations and suggestions for future research

This paper is a first exploration of the impact of gestural shifts, i.e., gestural omissions and/or additions, on the relation between primary participants' originals and of interpreters' renditions and thus on the information exchange process in healthcare interpreting. We provided a qualitative analysis of three excerpts taken from authentic interpreter-mediated medical consultations. However, future research could replicate the analyses presented here on a larger dataset and add a quantitative dimension—which we have not conducted in our analyses so far. In doing so, such a follow-up study could provide a stronger empirical basis and a more thorough understanding of the phenomena discussed here.

Following up on this, our dataset consisted of only patients with a Russian or a Turkish background. As we know that the use of gestures varies across cultures, it might be relevant to include more cultural and linguistic variation and zoom in on cross-cultural variation in the use of gesture. Finally, we have mainly looked at HCPs' utterances in the information exchange process. As this process also entails patients providing information about their lived experiences in order to make a diagnosis and in order to be able to participate the decision making, a follow-up study could replicate our analyses on their turns-at-talk and the interpretations thereof in order to corroborate our understanding of multimodal information exchange and interpretation strategies in healthcare communication.

6. References

Amato, A. (2004). The interpreter in multi-party medical encounters. In C. Wadensö, B. Englund Dimitrova & A-L. Nilsson (Eds.), *The critical link 4: professionalization of interpreting in the community. Selected papers from the 4th International Conference on Interpreting in Legal, Health and Social Service Settings. Stockholm, Sweden, 20-23 May 2004* (pp. 24-38). Benjamins.

Angelelli, C. (2019). *Healthcare interpreting explained*. Routledge.

Angelelli, C. (2004). *Medical interpreting and cross-cultural communication*. Cambridge University Press.

Aranguri, C., Davidson, B., & Ramirez, R. (2006). Patterns of Communication through Interpreters: A Detailed Sociolinguistic Analysis. *Journal of General Internal Medicine : JGIM*, *21*(6), 623–629. https://doi.org/10.1111/j.1525-1497.2006.00451.x

Beukeleers, I., Theys, L., Salaets, H., Krystallidou, D., Pype, P., & Brône, G. (2023). *Exploring the semiotic complexity of interpreting: how gestural shifts can impact the information exchange process*. Paper presented at the International Pragmatics Conference, July 2023, Brussels.

Beattie, G. & Shovelton, H. (2001). An experimental investigation of the role of different types of iconic gesture in communication. *Gesture*, *1*(2), 129–149. https://doi.org/10.1075/gest.1.2.03bea

Bolden, G. B. (2000). Toward understanding practices of medical interpreting: interpreters' involvement in history taking. *Discourse Studies*, *2*(4), 387–419. https://doi.org/10.1177/1461445600002004001

Cirillo, L. (2012). Managing affective communication in triadic exchanges: Interpreters' zero-renditions and non-renditions in doctor-patient talk. In C.J. Kellet Bidoli (Ed.), *Interpreting across Genres: Multiple Research Perspectives* (pp. 102-124). Trieste: EUT.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review 123*(3), 324-347. https://psycnet.apa.org/doi/10.1037/rev0000026

Clark, H. H. (1996). *Using language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Clark H.H. & Brennan, S.A. (1990). Grounding in communication. In L.B. Resnick, J.M. Levine & S.D. Teasly (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). American Psychological Association. https://doi.org/10.1037/10096-006

Davidson, B. (2000). The interpreter as institutional gatekeeper: The social-linguistic role of interpreters in Spanish-English medical discourse. *Journal of Sociolinguistics 4*(3), pp. 379-405. https://doi.org/10.1111/1467-9481.00121

Davitti, E. (2019). Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis. *Qualitative Research: QR*, *19*(1), 7–29. https://doi.org/10.1177/1468794118761492

De Wilde, J., Van Praet, E., & Van Vaerenbergh, Y. (2019). Language discordance and technological facilitation in health care service encounters. A contrastive experiment. In P. Garcés-Conejos Blitvich, L. Fernández-Amaya & M. de la O Hernández-López (Eds.), *Technology mediated service encounters (*p. 17-44). John Benjamins. https://doi.org/10.1075/pbns.300.01dew

Enfield, N. J. (2013). A "Composite Utterances" approach to meaning 45. Towards a grammar of gestures: A form-based view. In Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D. & Tessendorf, S. (Eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (pp. 202-217). De Gruyter. https://doi.org/10.1515/9783110261318.689

Enfield, N. (2009). *The anatomy of meaning: speech, gesture and composite utterances*. Cambridge University Press.

Gardner, R. (2001). *When listeners talk. Response tokens and listener stance.* John Benjamins Publishing Company.

Gerwing, J. & Allison, M. (2009). The relationship between verbal and gestural contributions in conversation: A comparison of three methods. *Gesture, 9*, 312–336. https://doi.org/10.1075/gest.9.3.03ger

Gutierrez, A. M., Statham, E. E., Robinson, J. O., Slashinski, M. J., Scollon, S., Bergstrom, K. L., Street, R. L., Parsons, D. W., Plon, S. E., & McGuire, A. L. (2019). Agents of empathy: How medical interpreters bridge sociocultural gaps in genomic sequencing disclosures with Spanish- speaking families. *Patient Education Counseling*, *102*(5), 895-901. https://doi.org/10.1016/j.pec.2018.12.012

Hsieh, E. (2016). *Bilingual health communication: Working with interpreters in cross-cultural care*. Routledge.

Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & Fröhlich, M. (Eds.), *Gesture and sign language in human-computer interaction* (pp. 23-35). Springer.

Major, G. & Napier, J. (2012). Interpreting and knowledge mediation in the healthcare setting: What do we really mean by "accuracy"? *Linguistica Antverpiensia, New Series – Themes in Translation Studies*. https://doi.org/10.52034/lanstts.v11i.30

McNeill D. (1992). *Hand and mind: What gesture reveals about thought*. University of Chicago Press.

Menichetti, J., Lie, H. C., Mellblom, A. V., Brembo, E. A., Eide, H., Gulbrandsen, P., Heyn, L., Saltveit, K. H., Strømme, H., Sundling, V., Turk, E., & Juvet, L. K. (2021). Tested communication strategies for providing information to patients in medical consultations: A scoping review and quality assessment of the literature. *Patient Education and Counseling*, *104*(8), 1891–1903. https://doi.org/10.1016/j.pec.2021.01.01

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, Cienki, A., Fricke, E., Ladewig, S., McNeill, D., & Teßendorf, S. (Eds.) *Body – Language – Communication. An international handbook on multimodality in human interaction* (pp. 1687-1702). De Gruyter. https://doi.org/10.1515/9783110302028.1687

Müller, C. (1998). *Redebegleitende Gesten: Kulturgeschichte – Theorie – Sprachvergleich*. Arno Spritz.

Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D. & Tessendorf, S. (Eds.). (2003). *Body – Language – Communication. An international handbook on multimodality in human interaction*. De Gruyter.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, *369*(1651), 20130296–20130296. https://doi.org/10.1098/rstb.2013.0296

Pöchhacker, F. (2022). *Introducing interpreting studies* (3rd ed.). Routledge.

Rowbotham, S., Holler, J., Lloyd, D., & Wearden, A. (2011). How do we communicate about pain? A systematic analysis of the semantic contribution of co-speech gestures in pain-focused conversations. *Journal of Nonverbal Behavior*, *36*(1), 1–21. https://doi.org/10.1007/s10919-011-0122-5

Schegloff, E. (2007). *Sequence organization in interaction: A primer in Conversation Analysis*. Cambridge University Press.

Schegloff, E. A. (2000). When "others" initiate repair. *Applied Linguistics*, *21*(2), 205–243. https://doi.org/10.1093/applin/21.2.205

Theys, L., Wermuth, C., Salaets, H., Pype, P., & Krystallidou, D. (2023). The co-construction of empathic communication in interpreter-mediated medical consultations: A qualitative analysis of interaction. *The International Journal of Translation and Interpreting Research*, *15*(1), 45–73. https://doi.org/10.12807/ti.115201.2023.a03

Theys, L. (2021). *The co-construction of empathic communication in interpreter-mediated medical consultations. An exploratory study using video recorded interactions and video stimulated recall interviews.* [Doctoral dissertation, KU Leuven, UGent].

Wadensjö, C. (1998). *Interpreting as interaction*. Longman. https://doi.org/10.4324/9781315842318

7. Acknowledgements

This study is part of the Right2Health-project funded by the NWO. This is a collaboration between the University of Amsterdam, University of Leuven, Utrecht University and the University of Ghent. We would like to thank the consortium for their feedback on earlier presentations about this study. We are also grateful to the two anonymous reviewers for their constructive feedback on previous versions of this manuscript. Finally, we would like to thank the patients, interpreters and healthcare providers to participate in the data collection and allow us to investigate language discordant communication in a natural setting.

8. Appendix

Transcription conventions

HCP	Healthcare professional
I	Interpreter
P	Patient
Speech	utterance as produced by the interlocutor
Speech	translation of the utterance into English
(.)	a brief pause (<0.2 ms)
(0.5)	duration of a pause in tenths of a second
[carcino]ma	start and end of overlapping speech
#Fig. 1	occurrence of gesture as illustrated in figure 1
----------	duration of gesture

Inez Beukeleers
University of Leuven, Department of Linguistics
Blijde-Inkomststraat 21
B-3000 Leuven
Belgium

Inez.beukeleers@kuleuven.be

Biography: Inez Beukeleers is a post-doctoral researcher and a lecturer in the Department of Linguistics at KU Leuven. Currently, she is working on the *Right2Health research project* (NWO) in which she investigates multimodal communication strategies in authentic language discordant medical consultations. She adopts a multifocal eye-tracking approach. The project zooms in on 1) how often healthcare providers and patients use visual information provision strategies and how this information is relayed by interpreters and 2) multimodal grounding strategies. The analysis compares communicative strategies that occur in consultations with professional interpreters with strategies that occur in consultations mediated by informal interpreters.

Laura Theys
Agentschap Integratie en Inburgering
Kanselarijstraat 17a
B-1000 Brussel
Belgium

laura.theys@integratie-inburgering.be

Biography: Laura Theys completed a PhD on empathic communication in public service interpreting in 2021. After that, she further specialised in public service interpreter training (KU Leuven, AgII & KTV Kennisnet) and project management (AgII). Currently, she is working as a project manager at AgII on the EU-WEBPSI project. This job entails designing a module on the organisation of certification tests for public service interpreters and organising the Belgian training for public service interpreters for languages of limited diffusion. In addition, she helps to coordinate and manage the training sessions in the partner countries.

Heidi Salaets
University of Leuven, Research Unit of Translation Studies
Sint-Jacobsmarkt 49-51
B-2000 Antwerpen
Belgium

Heidi.salaets@kuleuven.be

Biography: Heidi Salaets is an associate professor at the KU Leuven Faculty of Arts, Campus Antwerp. She mainly teaches in the Master of Interpreting program, covering subjects such as interpreting studies, ethics and professional conduct, note-taking techniques, interpreting techniques, community interpreting, and dialogue interpreting. Key areas of research are legal interpreting and interpreting in healthcare, which she approaches from an interdisciplinary perspective and also involving key stakeholders throughout the research phases. In order to generate a substantial impact of the research projects, Heidi Salaets offers interprofessional training to, among others, medical and legal staff.

Cornelia Wermuth
University of Leuven, Research Unit of Translation Studies
Sint-Jacobsmarkt 49-51
B-2000 Antwerpen
Belgium

Cornelia.wermuth@kuleuven.be

Biography: Cornelia Wermuth is associated professor at KU Leuven at the Department of Applied Linguistics. She lectures German grammar and Terminology and IT in the Bachelor in Applied Language Studies and Medical Translation in the Master in Translation. Her main areas of interest are specialized (medical) translation, terminology, translation tools and terminological computer applications, applied cognitive linguistics (medical sublanguage), and frame semantics. Her current research focuses on medical ontologies and their relationship to terminological concept systems.

Barbara Schouten
University of Amsterdam
Nieuwe Achtergracht 166
Post box 15791
1001 NG Amsterdam
The Netherlands

B.C.Schouten@uva.nl

Biography: Barbara Schouten's research interests focus on intercultural health communication, in particular on the role of language-related and culture-related factors explaining communication difficulties between healthcare providers and ethnic minority patients with low language proficiency in the host country's dominant language(s). In addition, she focuses on the use of technology in interventions to mitigate language- and culture-related barriers in communication.

Geert Brône
University of Leuven, Department of Linguistics
Blijde-Inkomststraat 21
B-3000 Leuven
Belgium

geert.brone@kuleuven.be

Biography: Geert Brône is full professor of German and general linguistics at the Linguistics Department of KU Leuven. An important methodological focus of his research in the past years has been the development of mobile eye-tracking technology as a tool for multimodal interaction analysis, a line of research that zooms in on the relation between verbal and nonverbal resources (including gesture, head movements, eye gaze and facial expressions) in human interaction.

Parallèles 37(1) April 2025

Sustaining embodied participation frameworks with gaze and head gesture in signed-to-spoken interpreting

Vibeke Bø

OsloMet – Oslo Metropolitan University

Abstract

This study investigates embodied participation frameworks in a signed-to-spoken interpreted encounter. Using a multimodal conversation analytical lens, the analysis demonstrate how interpreters exploit available semiotic resources to sustain participation frameworks. While participation frameworks are constantly negotiated in both same-language and interpreted interactions, this study puts forward unique challenges that arise in signed-to-spoken interpreted encounters: Although gaze is an important interactional resource, the nature of signed-to-spoken interpreting sometimes requires an alternative strategy because the gaze is occupied with perceiving the signed discourse. Head gestures have been found to serve as this alternative strategy. The notion of the coupled turn in interpreted encounters is supported, as it helps unravel these patterns that are unique to interpreted interaction. The naturalistic data in this study provide examples of how navigating two communicative needs simultaneously leads to several simultaneous processes of embodied conduct: The interpreter visually perceives and renders an utterance, while also interactionally indicating the addressee with a head gesture. Findings from this study highlight the need for further exploration of how interpreters navigate competing communicative demands. Moreover, signed-to-spoken interpreting exemplifies the diversity of language practices, pointing to the need for an inclusive approach to language practices.

Keywords

Simultaneous interpreting, head gesture, coupled turn, embodied participation framework, conversation analysis

 DOI:10.17462/para.2025.01.09

1. Introduction

Gaze is an important resource for negotiating participant roles in face-to-face interaction (Goffman, 1981; Kendon, 1967; Rossano, 2012; Stivers & Rossano, 2010). Moreover, in the complex participation frameworks of interpreted interaction, gaze is attributed an important function (e.g., Davitti & Pasquandrea, 2017; Napier, 2007; Vranjes & Bot, 2021; Wadensjö, 2017). Gaze and head gestures are described as interactional resources in a substantial body of literature on interpreted interaction in which two spoken languages are used in the interaction (Davitti, 2012; Mason, 2012; Pasquandrea, 2011; Vranjes et al., 2018; Vranjes & Brône, 2021). However, in signed-to-spoken interpreted interactions, the exploration of the role of gaze and head gestures is still in its early stages.

In this paper, I investigate one interpreter's gaze and head gestures used as interactional resources in signed-to-spoken interpreting, i.e., from Norwegian Sign Language (NTS) into Norwegian. As gaze is the only way of perceiving the signed utterance, I suggest there may be a "trade-off" (Vranjes & Brône, 2021) between the resources of gaze and head gestures in the context of conversational signed-to-spoken interpreting, i.e., head gestures are used instead of gaze. The interactional character of interpreting is now well established in the literature (E.g., Roy, 2000; Wadensjö, 1998). However, there are not many studies that consider interactional processes in terms of their semiotic characteristics. Moreover, in-depth analyses of conversational signed-to-spoken interpreting are almost absent from the field. The current qualitative study on a single case of informal naturalistic interpreted discourse can serve as a starting point in that respect as it offers some important insights into how interactional and semiotic resources are interconnected in the context of face-to-face dialogue interpreting.

This study foregrounds how the interpreter deploys the semiotic resources she has at her disposal in the situation. Applying the concept of the coupled turn (Poignant, 2021), I demonstrate how the interpreter puts considerable work into sustaining the embodied participation framework (Goffman, 1981; C. Goodwin, 2007), thus maintaining the interactional space (Mondada, 2007) between two different language ecologies.

2. Embodied participation framework

When people engage in any form of social interaction, they unavoidably gain the status of a participant (Goffman, 1981, 1986). The status affects relations and expectations within the interaction, such as "who is addressing whom, and who is supposed by whom to react how?" (Wadensjö, 2017, p. 127). The character of this participation is constituted by the norms of whatever activity they are engaged in (Rossano, 2012). Within the participation framework, speakers are ratified as such by co-participants (Goffman, 1986). This ratification may occur by means of different resources, among which gaze is considered especially important (C. Goodwin, 2007; M. H. Goodwin, 1980; Kendon, 1967, 1990; Rossano, 2012). The question arises as to what specific expectations and norms, considering embodied participation frameworks, are activated when the activity type is an interpreted encounter between a signed and a spoken language.

Participation status is not static but is organized moment-by-moment through at times subtle communication practices. Goodwin (2007) stresses how embodied participation frameworks can reveal "the interactive organization of action, and of the active work required to sustain it" (C. Goodwin, 2007, p. 63). Part of the process of organizing the embodied participation framework is positioning one's body to have "appropriate perceptual access to relevant phenomena" (C. Goodwin, 2007, p. 63). The notion of appropriate perceptual access lends itself perfectly to make sense of signed-to-spoken interpreted interaction, as the interpreter needs visual access to the signing participant in order to interpret. The framework helps us identify the intricate

patterns to sustain the embodied participation frameworks accomplished by the interpreter. In the following, I review literature on relevant spoken and signed interaction before reviewing studies on how participation frameworks are achieved in interpreted encounters.

2.1. Gaze and head movements in spoken and signed interaction

Gaze is an important resource for establishing types of participation in conversation, specifically for selecting the next speaker, as extensively documented in spoken language interaction (C. Goodwin, 1981; Heath, 1986; Kendon, 1967; Stivers & Rossano, 2010). Moreover, interlocutors frequently need to pay attention to something while simultaneously conducting an interactional project through gaze (Rossano, 2012). Goodwin (2007) describes how visibly orienting to both other participants and the environment results in a cooperative stance, demonstrating the joint accomplishment of the activity in progress. Further, the cooperative stance requires appropriate perceptual access, for which people need to position themselves as needed physically (C. Goodwin, 2007). The cooperative stance can be described in more local terms: Gaze and postural shift can allow participants to display recipiency and a body movement can elicit speech by the other participant, or it can elicit a gaze re-orientation (Heath, 1986). Head nods may contribute to topicalization in both spoken (Bernad-Mechó, 2017) and signed (Liddell, 1980; Sutton-Spence & Woll, 1999) discourse.

Concerning signed interaction, gaze behavior has been attributed several functions. Regarding participation frameworks, eye gaze plays a crucial role in seeking and yielding turns (Baker, 1977). In this seminal study on turn-taking in signed discourse, Baker (1977) investigates a small deaf meeting and finds that head nodding, combined with a palm-up gesture, functions as a means of claiming a turn. In informal conversations among deaf friends, Coates and Sutton-Spence (2001) found that participants mostly waited until eye contact was established before beginning their turn. In addition to contributing to participation frameworks, gaze plays an important role in organizing discourse. Janzen and colleagues (2023) compare two signed languages and find that gazing upwards represents something that is unknown or distant in time or place. In constructed dialogue sequences, gaze serves as a co-establishing resource (Young et al., 2012). Additionally, in signed discourse, gazing at a significant point in space attributes a specific meaning, which is especially exploited in highly depictive modes of discourse (Dudis, 2011; Roy, 2011). In these depictive discourse strategies, signing space is perceived as a stage on which discourse entities may be placed. This way of organizing the signing space involves real space blends in cognitive linguistic terms (Liddell, 2003) and may be described as indicative of the discourse complexity (Nilsson, 2023). In interpreted interaction, depicting strategies may pose additional cognitive challenges for interpreters because they may entail a transition of semiotic strategy in the interpreting process (Nilsson, 2010, 2023).

Importantly, the pattern of gaze behavior is dependent on activity type, i.e., gaze expectations by participants are associated with the ongoing course of action. Rossano (2012) stresses that there are different norms for gazing at co-interactants depending on conversational activity type, which supports the need to study gaze patterns in different types of interactions. However, these norms also depend on language-specific ecologies. In this study, interpreting will be conceptualized as a form of conversational activity type, thereby suggesting that it involves specific norm-governing gaze behavior. Thus, in addition to activity type, I suggest the semiotic character of the source utterance also affects the interpreter's gaze behavior. Moreover, an interpreted event is constituted by the presence of at least two language ecologies, and thus, two different sets of gaze behavior norms are represented in the same interactional encounter.

2.2. Embodied participation framework and gaze in interpreted interaction

In interpreted interaction the complexity of participation frameworks is increased compared

to monolingual conversations, as acknowledged by Wadensjö (1998) and Roy (2000). Consequently, the interpreter has specific professional conversational and communicative needs (Jucker et al., 2018; Vranjes & Bot, 2021). Gaze is ascribed several functions in spoken language interpreting, including turn taking (Hansen & Svennevig, 2021; Lang, 1978; Mason, 2012; Pasquandrea, 2011; Vranjes et al., 2018), sequence organization (Vranjes et al., 2018) and the signaling of position and epistemic authority (Davitti, 2012; Mason, 2012). The current study leans on findings from studies on the role of gaze and participation roles, resulting in a non-normative, interactionist, and dialogical view (Wadensjö, 2017). Furthermore, an interpreter's participation has been conceptualized in terms of a professional role-space (Llewellyn-Jones & Lee, 2014). Llewellyn-Jones and Lee (2014) claim that interpreted interaction depends on interactional signals from the interpreter. If the interpreter suppresses these signals, as some are trained to do, the interaction might be perceived as dysfunctional (Llewellyn-Jones & Lee, 2014, p. 39).

Some experimental studies have described linguistic phenomena of signed-to-spoken interpreting (Gabarró-López, 2024; Nilsson, 2010; Quinto-Pozos et al., 2015; Santiago et al., 2015). Nilsson (2010, p. 64), finds that the character of discourse affects interpreters' ability to render it appropriately. However, to investigate interactional resources, a real audience for renditions is required. One naturalistic study explored the teamwork between a deaf professional and two interpreters in the context of a formal, monological talk given by the deaf professional (Napier, 2007; Napier et al., 2008). Pause, nods, and eye contact were found as important discourse markers for achieving clarification and controlling the pace. Nodding (often co-occurring with a sign or gesture) was also found to serve the communicative function of reassuring that all was going well (Napier et al., 2008, p. 32). Also in a formal context, Henley and McKee (2020), using an interactional sociolinguistic approach, compared two interpreted meetings led by a deaf and a hearing chair-person. In the deaf-led meeting, they found gaze, nodding, and pointing to have important turn-allocation functions. They highlight the two sets of discourse norms present in a mixed meeting and find that only in the meeting led by a deaf chairperson were the visual discourse norms adhered to. This adherence was found to increase the perceived access by the deaf participants in the meeting (Henley & McKee, 2020).

Finally, one study investigates interpreted classroom group-work activities among deaf and hearing upper secondary school students. In this study the direction of interpreting is mostly spoken-to-signed, as the deaf student is rarely ratified as a member of the hearing students (Berge, 2018, p. 108). The few instances described of signed-to-spoken renderings are consequences of the interpreter's negotiation of the participation status of the deaf student by means of, e.g., gaze and gestures (Berge, 2018, p. 108). In signed-to-spoken renditions, the interpreter exploits body leans and eye contact to indicate the addressee(s) of the signed utterances (Berge, 2018).

2.3. Interactional space and the coupled turn

In Conversation Analysis (CA), the adjacency pair consists of two self-contained turn-construction units (Sacks et al., 1974). The organization of interpreted interaction involves a turn-construction unit that is not self-contained, i.e., the interpreter's contribution is better viewed as the extension of the original utterance than an independent turn. This has led to the notion of a coupled turn, consisting of the original utterance and its rendition (Poignant, 2021). The notion of the coupled turn helps to understand how the interpreter manages to create a domain of conversation (Ciolek & Kendon, 1980, p. 237) or a joint interactional space (Mondada, 2013) by means of embodied resources. Sometimes, the interactional space of an interpreted encounter needs extra work to be negotiated according to ratified participation

roles. It is not always the case that hearing participants who are not used to interpreted interaction look at the deaf participant holding the floor. As the interpreter is making a signed utterance audible, people tend to look at the interpreter, while the principal is actually a deaf participant (Napier et al., 2019).

The research question of the present study is: How does the interpreter accomplish and sustain the embodied participation framework in a signed-to-spoken interpreted conversation? In particular, I focus on the role of gaze and head gestures and the intricate pattern of their interdependency. Findings may serve as evidence of the notion of a coupled turn in interpreted interaction.

3. Data and method

The analysis is based on naturalistic data consisting of one video-recorded informal lunch conversation (duration: 42:35 min) with two deaf participants, one non-signing hearing participant and an interpreter. The deaf participants and the interpreter have been colleagues for many years, familiar with each other. The interpreter is trained (in Norway, interpreters are required to have a BA to be qualified) and has more than ten years of experience. The hearing participant works in the same corridor, but in a different department and was previously unacquainted with both participants. She had very limited to no knowledge of sign language or deaf people in general and became intrigued by the subject. This situation led the conversation to focus on being and growing up deaf. This theme proved to benefit the research focus because many utterances were directed from the deaf participants to the hearing participant. The conversation was initiated by me, approaching participants by email (in Norwegian). The data is thus co-constituted between the researcher and participants in the study (Mondada, 2006). However, the participants are actual colleagues and are in a situation where there is something real at stake; they remain in a common workspace after this conversation.

The conversation was video recorded with two cameras while I was present in the room to manage the recording. The choice to stay in the room could be perceived as unfortunate because of my position in the field as an interpreter, interpreter trainer, and interpreting researcher. For this reason, I ensured that the participating interpreter had not been my student. The possibility of affecting the ongoing course of interaction in some way is nevertheless difficult to entirely dismiss; a human presence will always affect the room. The decision was also affected by the availability of data, as technical issues could compromise the quality of the recordings. In several instances, adjustments to camera angles were required due to participants changing their positions.

The study was granted approval by the Norwegian Centre for Research Data (SIKT). All participants signed written informed consent forms, which stated how the data would be used and presented. All participants agreed to openly sharing the data, without anonymization. Even though I have been granted permission to publish pictures and video clips, this does not relieve the researcher from treating participants as carefully and responsibly as possible (Skedsmo, 2021, p. 83). Thus, the names of the participants are changed to pseudonyms, following the alphabet: Anna, Beatrice (deaf participants) and Cora (hearing non-signing participant).

The data was annotated in ELAN (Crasborn & Sloetjes, 2008). The videos from the two cameras were aligned to display them in the same ELAN file. Initially, I identified all instances in which the interpreter orients towards the hearing participant with a head gesture, with or without gaze. Next, the corresponding NTS source utterance and gaze direction of the signing participant were annotated on two separate tiers. This was done in order to see how the gaze patterns of a rendition aligned with the original utterance in a coupled turn. Finally, the interpreter's verbal rendition (orthographically transcribed), gaze direction, and head gestures

were annotated on three separate tiers. The annotations of head gestures of spoken language renditions are inspired by the MUMIN schema guidelines (Allwood et al., 2007). The approach in this qualitative study is informed by multimodal conversation analysis (C. Goodwin, 2000; Mondada, 2014; Deppermann & Streeck, 2018).

To represent both the signed and spoken discourse of this interpreted event, I consulted different transcription traditions and developed an annotation guide in accordance with the research focus of the current study. The annotations of NTS discourse in this study are highly influenced by the guidelines used for Auslan[1] (Johnston, 2019). Since NTS does not have a written form, the annotations follow the tradition of glossing, which entails denoting each sign an English word that is close in meaning, written in SMALL CAPS. A gloss is not a translation but a lemma to represent signed discourse in written form. Importantly, though widespread in the field, this tradition is problematic because of the risk of signed languages being represented as a simple version of a spoken language (Janzen & Shaffer, 2023; Rosenthal, 2009). In this paper, the annotations of signed discourse are minimalistic, and readers are encouraged to view video clips to see the signed source utterance analyzed. The three short sequences analyzed for this paper can be found here: https://osf.io/n4c79/?view_only=e7af211d65c5485787a5848f0f196a7a.

The multimodal transcription conventions of embodied conduct are highly influenced by Mondada (2018). The full list of annotation conventions can be found in the Appendix. Still images from the open dataset are provided in annotations.

4. Results and analysis

This section includes the analysis of four extracts from the conversational data in which the addressee is the hearing participant. Overall, a total of 174 sequences were identified in which the interpreter gazes or moves her head (or both) towards the hearing participant while interpreting from NTS to Norwegian. The examples shown in this paper are selected to represent indicative behavior with and without gaze, and also to show a variety of head gestures. In most cases, the head gesture movement consists of a mixture of tilting (governed by the top of the head) and turning (governed by the chin). This variation may be explained by considering seating arrangements, displayed in figure 1.

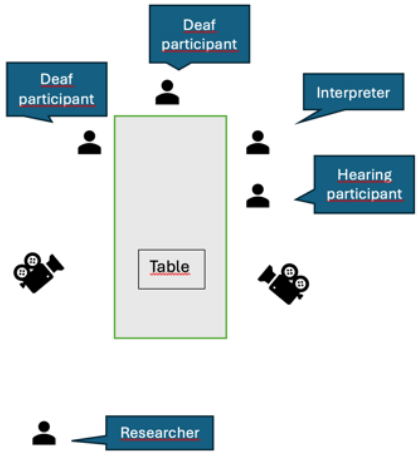


Figure 1. Seating arrangements

Consequently, all variation of head gestures without gaze is treated under the umbrella of head gestures. The types of head gestures are illustrated in figures 2-5. Figure 2 displays one of the examples where the interpreter includes gaze to indicate the addressee is provided:

[1] The majority signed language used in Australia.

Figure 2. Head gesture (side-turn) and gaze

In figures 3-5, I present examples of the interpreter indicating the addressee without gaze, realized with different head gestures. However, the positions are all in some way oriented towards the hearing participant. Examples of three different realizations without gaze are provided in the following:

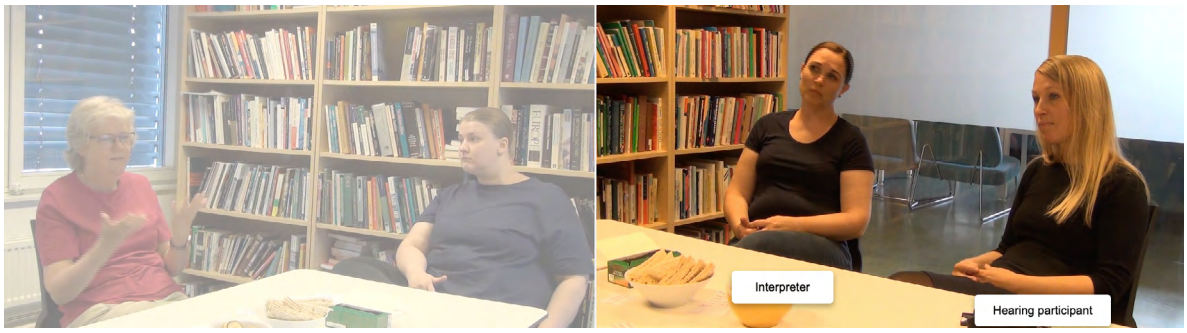


Figure 3. Head gesture: side-tilt

Figure 4. Head gesture: side-turn

Figure 5. Head gesture: back

The difference in form was not found to reflect a difference in meaning but is nevertheless presented to potentially serve as a starting point for future studies.

To get an impression of the relative frequency and distribution of the interactional resource of head gestures with and without gaze, see Table 1:

Gaze and head gestures used to visually indicate the addressee in spoken Norwegian renditions	Number of occurrences
Head gesture (side-turn) with gaze	24
Head gesture without gaze	150
Total	174

Table 1. Distribution of gaze direction and head gesture

From Table 1 we can see a total of n=174 tokens of rendered utterances (in spoken Norwegian) indicating the hearing participant (named Cora) as addressee with visual resources, with or without gaze. There is quite a small category (n=24) in which the interpreter directs her gaze towards the hearing participant. These instances consistently co-occur with a side-turn head gesture. The larger second category (n=150) are instances without gaze in which there is a variety of combinations of head-tilt and head-turn gestures. While it would be possible to categorize this further, according to type of head gesture, I leave more fine-grained analytic work concerning head gestures to future studies with a larger body of data. For this study, the point is to show how the embodied participation framework is navigated and affected by the signals of participation and the semiotic character of the source utterance.

In what follows, in-depth multimodal conversation analysis of four extracts from the recorded conversation are provided. The first two examples represent examples in which the interpreter shifts her gaze, joining the gaze direction of her co-participant in a coupled turn.

4.1. Interactional space and the coupled turn

In the dataset, head gestures frequently accompany an utterance directed towards the hearing participant but are not necessarily followed by gazing in the same direction. The gaze pattern I analyze in Extract 1 only occurs in n=24 instances from the data (see Table 1). In this extract, the interpreter's gaze is briefly directed towards the hearing participant, subsequent to a wh-question[2]. Interactionally, the interpreter selects the next speaker, reflecting the gaze behavior accompanying the source utterance, which indicates Cora (the hearing participant) selected as next speaker.

After figuring out that they have never met, but that their offices are actually quite close to each other, Anna asks Cora how long she has worked there.

[2] In Extract 1, there is a small particle *a* (line 2), marked <DM:Q>, and it is thus labelled a discourse marker. This is not a regular question word; it is a Norwegian way of signaling the request of a response in an informal style.

Extract 1:

1. Anna	GLOSS	#HOW-LONG [BE-AT #PRO-2p #COME-HERE
	Gaze	*Cora----------->>
	Trans	How long have you been here?
2. INT	Norwegian	# [men hvor #lenge har du #vært her a
	Head G	*Neutral---* *side-turn----------*
	Gaze	*Anna----------* *Cora(0.3)-* *Anna>>
	Trans	<DM:but>how long have you been here <DM:Q>

Images **#1.1** **#1.2** **#1.3**

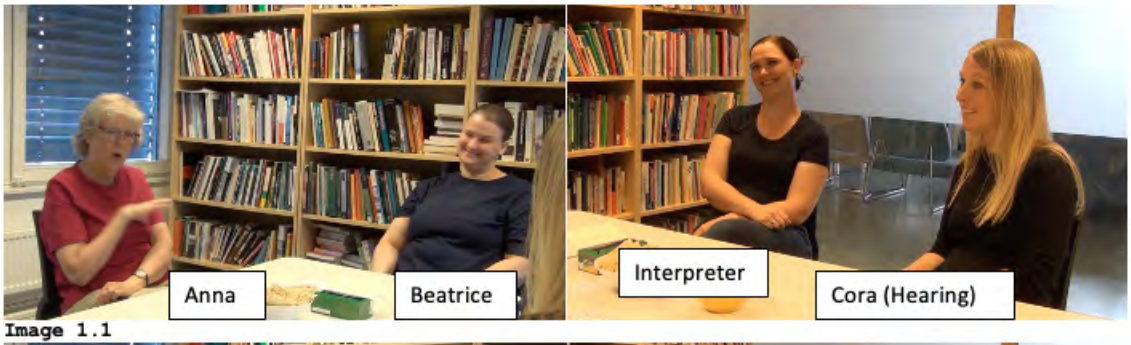


Image 1.1

Image 1.2

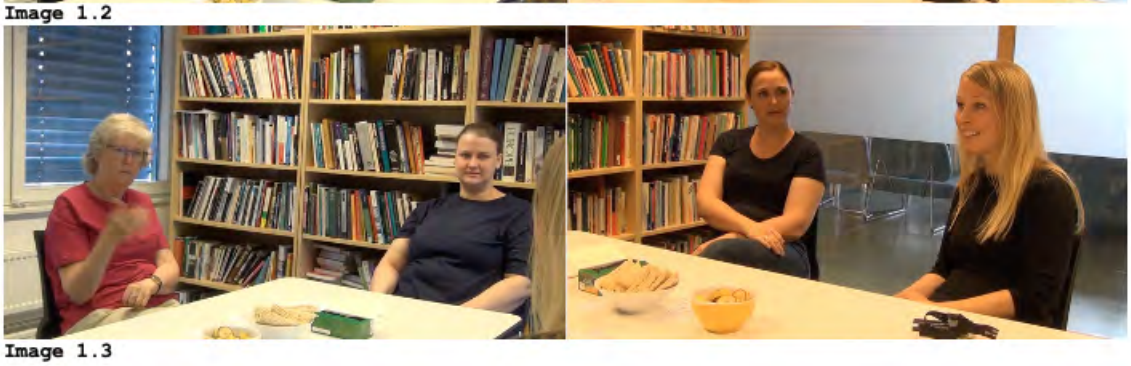


Image 1.3

Anna self-selects and produces a wh-question while gazing (and pointing) at Cora, who is thus selected as the next speaker: (´how long have you been here?´) (line 1). Anna maintains her gaze on the addressee while asking the question, aligning with observations made in monolingual encounters concerning gaze accompanying questions (Rossano, 2010; Stivers et al., 2009).

Before producing a rendition, the interpreter's head is slightly tilted towards her right (line 2, image 1.1). Starting the rendition, her head immediately turns slightly towards the hearing participant (line 2, image 1.2), signaling her emerging turn. The rendition includes a deictic pronoun 'you', a recipient indicator (Lerner, 2003, p. 182), making the next speaker explicitly addressed. This pronoun is immediately followed by an embodied orientation towards the hearing participant by a gaze shift (line 2). All participants now share a mutual focus of attention with their gaze (image 1.3), establishing Cora as the next speaker in the embodied

participation framework. The interpreter's gaze towards the hearing participant (0.3 sec) is not reciprocated, nor the interpreter's gaze back to the deaf participant. The interpreter is thus not treated as speaker in the participation framework. The hearing participant has her gaze directed towards Anna throughout the sequence, which is evidence that Anna is treated as ratified speaker (Goffman, 1981). This is not always the case in signed-to-spoken interpreting as hearing interlocutors tend to look towards the interpreter instead of the deaf signer (Napier et al., 2019). The interpreter's gaze is held for 0.3 seconds (image 1.3) before returning to Anna. The interpreter receives no gaze throughout the sequence. The absence of orientation towards the interpreter's gaze supports the notion of the coupled turn (Poignant, 2021), shared between Anna and the interpreter. The choice of prioritizing a gaze shift in a context with potential signed utterances that the interpreter needs to monitor, speaks to gaze as a powerful signal for selecting the next speaker (C. Goodwin, 2007; Rossano, 2012).

This example is illustrative of how the interpreter, when rendering a direct question accompanied by a gaze behavior selecting the next speaker, may join the gaze behavior of the signing participant and thus conduct a full shift of gaze, despite her perceptual requirement to look at a signing participant. Copying the gaze behavior of the speaker may serve as evidence for the coupled turn in interpreted interaction, as both participants (ratified speaker and the interpreter) cooperate in selecting the next speaker with gaze. We now move on to another example in which the gaze behavior in the original utterance is more complex, and where the physical angles of seating arrangements add to the complexity. To demonstrate how the interpreter organizes her gaze behavior in a rendition when faced with a more complex gaze pattern in the original utterance, the following example represents an instance with a very swift gaze shift in the rendition, reflecting a more indecisive gaze pattern from the original utterance.

In Extract 2, Beatrice has just explained that growing up, she sometimes had speech therapy, like most deaf children (in Norway), and she did not like it. The utterance, towards the end of her turn, summarizes that she is pleased that period of her life is finished. In the previous example, where we saw a direct question, Anna consistently gazed towards the hearing participant throughout the utterance. Beatrice displays a different gaze pattern: after a short gaze towards the hearing participant, she shifts her gaze towards a high location in signing space (gazing upwards). An upwards gaze is used to signal "distance" in time and/or space (Janzen et al., 2023), which aligns with the pattern observed in this example: Beatrice addresses the distant past in describing experiences from her childhood. Note how the interpreter's head position is high, possibly orienting towards the same area as Beatrice's gaze (image 2.1). The specific gaze behavior of Beatrice is annotated in Extract 2 (see lines 1 and 3). When producing her rendition in the coupled turn, the gaze pattern of the interpreter can be seen in the images 2.1-2.4. In image 2.2, observe how the interpreter orients towards the hearing participant with a gaze shift.

Extract 2:

1. Beatrice	GLOSS	[#FEEL HAPPY FINISHED PT-DET PERIOD
		*folding hands--->
	Gaze	*Cora* *Up----------* *INT/Cora(0.7)->
	Trans	I am glad to be finished with that period
2. INT	NOR	[#<DM:så> <DM:ja> jeg er egentlig veldig glad
	Head G	*Back---------->
	Gaze	*Beatrice---------->
	Trans	<DM:so><DM:yes> I am actually very happy
3. Cora	Gaze	*#Beatrice---------->
	Images	**#2.1**

Image 2.1

4. Beatrice	Gaze	INT/Cora----------**Anna(1.1)** #INT/Cora(0.4)**Anna#(0.6)>>
		folding hands--->>
5. INT	NOR	for at jeg er ferdig med #den #perioden
	Head G	Back--------* *side-turn----------*
	Gaze	Beatrice----* *Cora----------* *Beatrice->>
	Trans	for being finished with that period
6. Cora	Gaze	*Beatrice----------#----------* #*INT(0.7)-*
	Images	**#2.2** **#2.3**

Image 2.2

Image 2.3

Beatrice signals the potential completion of her turn placing her hands on her lap (lines 1 and 3). However, her facial expression also signals marking of a stance (pursed lips; images 2.2-2.3) toward her own story (Ruusuvuori & Peräkylä, 2009, p. 386). The pursed, smiling lips may signal contentment because she is finished with the period of her life that included speech therapy, but it may also mark decisiveness, i.e., something she feels strongly about. Simultaneously, she displays a pattern of gaze behavior co-occurring with her facial expression. Accompanying her hands on her lap, her gaze is directed towards the interpreter (0.7 sec.), who is in the middle of her rendition. This aligns with the monitoring reported from other deaf professionals working with interpreters (Haug et al., 2017). Then, she looks at Anna (1.1 sec.), opening a possibility for her to take the floor. As Anna does not take the floor, Beatrice returns her gaze to the interpreter (0.4 sec.), and finally to Cora (0.6 sec.) (line 3). In sum, this sequence of gaze behavior with a facial expression of stance (Feyaerts et al., 2022; Ruusuvuori & Peräkylä, 2009) lasts 2.8 seconds. Interactionally, she signals readiness to yield the floor to someone else. Note however, that she does not select the next speaker. The gaze behavior, where she looks at all participants in turn (including the interpreter), leaves the floor to a potential self-selected speaker.

In the rendition part of the coupled turn, the interpreter is not provided with sufficient embodied cues to treat anyone as the selected next speaker, and her gaze toward Cora is very brief before returning to Beatrice. However, note that the interpreter's gaze toward Cora is reciprocated (image 2.3, indicated with arrow) which again speaks to the power of gaze in conversation in general. While a gaze would normally be evidence that the interpreter is treated as a speaker and thus contests the participation framework, Cora gives several signals that she treats Beatrice as the speaker. She orients towards Beatrice with gaze after 0.7 seconds. In the previous example, the absence of gaze served as evidence that the interpreter was treated as different than the other participants, but still considered an active participant in the embodied participation framework. This may serve as evidence for the notion of the coupled turn, considering that Cora's gaze toward the interpreter is visible for Anna. This suggests that Cora, by gazing at both the interpreter and Anna, sequentially acknowledges the coupled turn and thus signals a cooperative stance (C. Goodwin, 2007) towards the participation framework.

Given the seating arrangements of this situation, the interpreter's orientation towards Beatrice leaves Cora almost behind her, outside of her visually accessible area, making the interactional space between them physically different. This may be the reason why the head gesture is more tilted backwards than in the previous example (image 2.1, line 2-3) as this will increase her physical peripheral range of vision. The moment Beatrice places her hands in her lap, signaling readiness to yield the floor, the interpreter initiates her shift of gaze almost simultaneously (0.1 seconds subsequent of placing the hands in her lap). Thus, the interpreter's initiated gaze shift occurs immediately after Beatrice is orienting towards the interpreter with her gaze. Gaze in this sequence is timed as if the interpreter is forwarding the gaze to Cora (see timing of this gaze behavior in lines 3 and 4). As we have already seen, the gaze is only a very quick orientation towards Cora before returning her gaze back to Beatrice. In addition to semiotic work relevant to the embodied participation framework, seating arrangements may also impact this pattern: The seating angle now leaves Beatrice outside of the interpreter's visually accessible space, making the interpreter unable to monitor and recognize communication signals.

Summing up, this example illustrates how the interpreter restricts her gaze behavior to align with the interactional signals of the original utterance. In addition, it demonstrates how seating arrangements in interpreted interactions may impact the possibilities of maintaining the interactional space. The interpreter initiates a gaze shift but does not fully direct it toward Cora (see Image 2.2); instead, she immediately returns her gaze to the other participants. This

may in part be due to the gaze pattern in the original utterance, where Beatrice shifts her gaze between all participants in the interaction, not selecting a next speaker. However, this swift gaze was sufficient for Cora to reciprocate it, possibly displaying a cooperative stance towards the participation framework. Moreover, depending on visual access because the interpreter does not know where, or in which modality the next utterance will come from, she needs to position herself to have visual access (C. Goodwin, 2007) to the deaf signing participants in particular. Having seen two examples in which the interpreter shifts her gaze to indicate the direction of utterances, we will see instances in the following two examples in which gaze is not shifted; the interactional semiotic work is conducted by other resources, specifically head gesture.

4.2. Head gestures without gaze

The data presented above has demonstrated how the interpreter indicates the addressee using both gaze and head gesture. However, most of the indicative behavior towards the addressee in the data occurs without gaze and thus solely with a head gesture. The next two examples are originally one sequence, divided into two extracts because the interpreter displays two different head gestures in the sequence. The semiotic work to sustain the embodied participation framework is subtle, but significant, as it represents a pervasive pattern of the interpreter's embodied interactional resources (see Table 1).

Extract 3 depicts an example of indicative behavior solely with head gesture. Anna talks about a dog she used to have. As the dog was also deaf, they would both be startled if a car came up behind them. Anna provides Cora with some information, that the dog was also deaf. She selects Cora as the addressee with her consistent gaze throughout this piece of information. However, Anna does not signal any readiness to leave the floor to someone else, as she gazes towards the signing space in the next sequence (line 1). Early in her rendition, the interpreter makes a slight head gesture towards Cora (indicated with an arrow). Consider the difference in head position between images 3.1 and 3.2:

Extract 3:

1. Anna	GLOSS	POSS-1p.s	#DOG ALSO	[DEAF PRO-1P.s#DEAF
	Gaze	*Cora----------		*SS----------
	Trans	My dog is also deaf		
2. INT	NOR		#	[eh (.)og #hunden min var også døv
	Head G			*side-turn/downward--*
	Gaze	*Anna----------		---------->
	Trans			eh and my dog was also deaf
	Images		**#3.1**	**#3.2**

Image 3.1

Image 3.2

In this example, the interpreters' slightly downward side-tilt (image 3.2) accompanies the introduction of a new topic: the dog (recently introduced) is deaf, as its owner. Anna topicalizes the sign DEAF with a head nod (Liddell, 1980; Sutton-Spence & Woll, 1999). This topicalization of a discourse entity is also reflected in the rendition. It may be an instance of copying behavior, although head nods are also used for topicalization in spoken discourse (Bernad-Mechó, 2017). However, the direction in which the head is directed simultaneously signals visual orientation towards the hearing participant, semiotically indicating the addressee in a coupled turn. This argument is further supported if we consider image 3.2 in extract 3 in which we can see that the head position is moved out of the optimal position at which she can look straight ahead to perceive the signed utterance she is interpreting. The communicational need to perceive what is signed, and the communicational need of interactionally indicating the direction of an utterance compete for the resource of gaze. In alternative terms, there is a "trade-off" (Vranjes & Brône, 2021) between gaze and head gesture. The interpreter's head gesture oriented towards the addressee creates a joint orientation towards the hearing participant, which again reveals the cooperation between the interpreter and Anna in the coupled turn.

In the continuation of this sequence (see Extract 4), the gaze behavior of Anna is more varied. Moreover, the semiotic character of discourse also has consequences for the interactional space in this example. Consider again the difference between head positions in images 4.1 and 4.2.

Extract 4:

1. Anna	GLOSS	CANNOT HEAR #[BEHIND CAR DEP:VEHICLE-PASSING #DEP:HOLD-LEASH
	Gaze	*SS---* *Co* *SS--* *Co* *INT---* *Cora----* *SS----------*
	Trans	If a car came behind us, none of us could hear it
2. INT	NOR	#[så så vi hørte(.)ingen#av oss hørte
	Head G	*side-tilt---------->
	Gaze	*Anna---------->
	Trans	so, so we didn't hear, neither of us heard

Images #4.1 #4.2

Image 4.1

Image 4.2

3. INT	NOR	det kom bil bak <DM:da>
	Head G	side-tilt--*
	Gaze	Anna---------->>
	Trans	a car came up behind us <DM:then>

Anna is engaged in a discourse semiotically characterized by depiction (Dingemanse, 2015; Ferrara & Hodge, 2018), in which she depicts a car coming up from behind (her right hand represents the car), and how she and her dog would react if that happened (enacting how she would hold the dog leash). The depictive sign system prompts recipients to imagine what the depicted entities look like (Clark, 1996; Dingemanse, 2015; Ferrara & Hodge, 2018). This discourse mechanism organizes the interactional space in a specific way: the space in front of Anna is now perceived as a stage for the invisible referred events to unfold. This highly depictive discourse will in turn have its effect on the perceived interactional space, and could affect the interpreter's cognitive load (Nilsson, 2010). Depictive sequences like this are typically organized in part by gaze: the signer establishes specific areas in signing space as significant by looking at them (Dudis, 2011; Roy, 2011; Young et al., 2012). This is also the case here: Anna shifts her gaze between signing space, Cora (the hearing participant) and the interpreter (line 1). Anna's gaze pattern is not reflected by the interpreter, whose gaze is not shifted, but consistently directed towards Anna in the coupled turn with a Norwegian rendition. However, the embodied conduct of leaning towards Cora with a side-tilt head gesture allows her to visually indicate the addressee without gaze. This indicative behavior is in part an independent choice by the interpreter, as Anna is shifting her gaze between participants and signing space. Note that the interpreter's head gesture is somewhat elevated, which may reflect the interpreter's need to obtain a bird's eye view of the interactional space, as Anna is actively exploiting the 3D possibilities of depictive signed discourse. The highly depictive character of discourse is also physically affecting the interactional space of this conversation.

5. Discussion

The interpreter is faced with the complicated task of being highly attentive towards the deaf participant in order to perceive the NTS utterance, while simultaneously including the hearing participant in the interaction. If this interactional goal is not attained, the perceived participation framework is at stake (Llewellyn-Jones & Lee, 2014, p. 39). When the interpreter's attentiveness towards the deaf participant requires gaze, she needs to make use of other available resources, as demonstrated in the current study. The findings support previous claims of gaze and head movements in signed-to-spoken interaction, i.e., they are important interactional resources (Henley & McKee, 2020; Napier et al., 2008), and the character of discourse may affect the interpreter's language practices (Nilsson, 2010). This study complements the literature on signed-to-spoken interpreting with conversational data.

Rossano (2012, p. 313) argues that earlier studies on gaze behavior have not accounted for the different expectations and norms of gaze behavior of different activity types. Regarding interpreting as an activity type, this study supports this view. Moreover, based on observations from the last example (see section 4.2), it was argued that the interactional space is not only affected by activity type, but also by the semiotic character of discourse. Discourse in this example was characterized by depiction as a semiotic strategy. Consequently, part of the physical space between interlocutors was conceptualized as a scene in which discourse entities were placed, in part by looking at these places. When gaze has this additional semiotic function, a different expectation regarding gaze behavior emerges, in turn affecting the interactional space. Thus, I argue we should not only account for activity type, but also the semiotic character of discourse when discussing participation frameworks and gaze. The semiotic lens applied to gaze behavior also highlights that two sets of discourse norms are present in one interactional event (Henley & McKee, 2020), or a "constant overlap between target and source environment" (Wadensjö, 2004, p. 105) which is constitutive of the face-to-face interpreted event. Bringing this feature of the interpreting task to the forefront has implications for how we discuss the task of interpreting.

Interpreters have been found to inhabit a key coordinating role in interaction in dialogue interpreting settings (e.g., Mason, 2012; Pasquandrea, 2011; Wadensjö, 1998). While this "key coordinating role" could be conceptualized as an interpreter-specific behavior, coordinating discourse is in fact a fundamental characteristic of interaction in general, recognized since Goffman (1963, 1986), and emerges as a consequence of the interpreter being an active participant in interaction. This study has provided examples of how an interpreter can find herself with competing needs between her conversational needs and the role of coordinating discourse as an interpreter (Vranjes & Bot, 2021). This is specifically observed in the need to perceive an utterance while visually indicating the addressee of the utterance. The need to focus on the signed discourse might be affected by the semiotic character of the utterance: Highly depictive sequences may pose specifically demanding cognitive tasks for the interpreter (Nilsson, 2010, 2023).

When the resource of gaze is occupied with perception, we have seen examples where resources are organized successively, i.e., the interpreter shifts her gaze towards the addressee (the hearing participant) after perceiving the signed utterance. However, the majority of instances in which the addressee is indicated visually occur without gaze. In these instances, head gestures have the interactional task of sustaining the participation framework, and thus the interactional space. Moreover, due to the different positions of the two deaf participants, the interactional spaces provide different possibilities to shift the gaze, as the interpreter risks losing visual access due to her perceptual capacity. Thus, she positions herself with different head positions to ensure visual access. The notion of an interactional space foregrounds what is at stake: shifting the gaze might entail losing the common interactional space. Thus, the interpreter finds strategies of accommodating space to her communicative needs (Jucker et al., 2018, p. 99).

6. Conclusions

In this study, I have demonstrated how the embodied participation frameworks of one signed-to-spoken interpreted encounter are constantly negotiated with intricate patterns of semiotic resources, similar to the patterns of participation frameworks in general (Goffman, 1981, 1986; C. Goodwin, 2007). However, there are some specific ecological factors of these situations that will inevitably affect how participation frameworks are accomplished. First, the interpreter's gaze is consistently occupied with perceiving the NTS utterance, which results in the constant navigating of (at least) two simultaneous communicational needs: perception of signed discourse and indicating the addressee of renditions. In the first two examples, the interpreter nevertheless prioritized a gaze shift, which speaks to gaze as a powerful resource of indicating the addressee of an utterance (C. Goodwin, 2007; Rossano, 2012). In the last two examples, representing the majority of instances in which the interpreter visually oriented towards the addressee, there was no gaze shift involved, only head gestures.

Applying the notion of the coupled turn (Poignant, 2021) I have demonstrated how dialogue interpreting requires a specific form of collaboration between all parties involved: When the speaker selects the addressee of an utterance by means of gaze, the interpreter may reflect this gaze direction. If the gaze is occupied with perception, the interpreter may instead exploit head gestures to visually mark the addressee of the rendition. Thus, the interpreter navigates two simultaneous interactional processes, perceiving an NTS utterance on the one hand and producing a spoken language utterance on the other.

This study is limited in terms of the size of data, and further investigations are needed to explore gaze and head gesture patterns of different constellations of participants. Furthermore, this study only considers gaze and head gestures; in future studies, a larger variety of visual

resources could be investigated, e.g., manual gestures, other facial expressions and body leans. Also, this study only briefly looks at the involvement from the hearing participant. To learn more about the intricate patterns of signed-to-spoken interpreting, more focus should be directed towards the hearing participants of such encounters.

Finally, I claim the methodological tools of multimodal conversation analysis have proven useful to highlight the organization of resources deployed to sustain the embodied participation frameworks in interpreted discourse. The framework has allowed for the scrutiny of the semiotic characteristics of resources at play, which is useful to increase specificity in terminology when discussing language practices of interpreters. Also, it is a framework that is not concerned with the vehicle of a semiotic resource, or its linguistic status, which makes it a more inclusive approach. The explorations of interpreted discourse in this qualitative study add to our knowledge regarding semiotic strategies deployed interactionally in a signed-to-spoken interpreting context. The claim in this paper is that this approach is useful when documenting and analyzing the language practices of interpreters as it foregrounds that interactional and pragmatic resources are crucial parts of an interpreter's competence.

7. Acknowledgements

First and foremost, I am grateful to the deaf and hearing participants for their willingness to be published without anonymization. I also extend my sincere appreciation to the interpreter for openly sharing her professional practice. Additionally, I am deeply thankful to my supervisors, Professor H. Haualand and Professor A.-L. Nilsson, as well as Associate Professor J.P.B. Hansen, for their invaluable support in shaping this article. Lastly, I express my gratitude to the two anonymous reviewers for their insightful comments, which have significantly improved this work.

8. References

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., & Paggio, P. (2007). The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation*, *41*(3), 273–287. https://doi.org/10.1007/s10579-007-9061-5

Baker, C. (1977). Regulators and turn-taking in American Sign Language discourse. In L. A. Friedman (Ed.), *On the other hand: New perspectives on American Sign Language.* (pp. 215–236). Academic Press, INC.

Berge, S. S. (2018). How sign language interpreters use multimodal actions to coordinate turn-taking in group work between deaf and hearing upper secondary school students. *Interpreting: International Journal of Research and Practice in Interpreting*, *20*(1), 96–125. https://doi.org/10.1075/intp.00004.ber

Bernad-Mechó, E. (2017). Metadiscourse and Topic Introductions in an Academic Lecture: A Multimodal Insight. *Multimodal Communication*, *6*(1), 39–60. https://doi.org/10.1515/mc-2016-0030

Ciolek, T. M., & Kendon, A. (1980). Environment and the Spatial Arrangement of Conversational Encounters. *Sociological Inquiry*, *50*(3–4), 237–271. https://doi.org/10.1111/j.1475-682X.1980.tb00022.x

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Coates, J., & Sutton-Spence, R. (2001). Turn-taking patterns in deaf conversation. *Journal of Sociolinguistics*, *5*(4), 507–529. https://doi.org/10.1111/1467-9481.00162

Crasborn, O., & Sloetjes, H. (2008). Enhanced ELAN functionality for sign language corpora. In O. Crasborn, E. Efthimiou, T. Hanke, E. D. Thoutenhoofd, & I. Zwitserlood (Eds.), *6th International Conference on Language Resources and Evaluation: Third workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora, 26 May-1 June 2008, Marrakech, Morocco* (pp. 39–43). European Language Resources Association.

Davitti, E. (2012). *Dialogue Interpreting as Intercultural Mediation Integrating Talk and Gaze in the Analysis of Mediated Parent-Teacher Meetings*. ProQuest Dissertations Publishing.

Davitti, E., & Pasquandrea, S. (2017). Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings. *Journal of Pragmatics*, *107*, 105–128. https://doi.org/10.1016/j.pragma.2016.04.008

Deppermann, A., & Streeck, J. (2018). *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (Vol. 293). John Benjamins Publishing Company.

Dingemanse, M. (2015). Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language*, *39*(4), 946–970. https://doi.org/10.1075/sl.39.4.05din

Dudis, P. (2011). The body in scene depictions. In C. B. Roy (Ed.), *Discourse in signed languages* (Vol. 17, pp. 3–45). Gallaudet University Press.

Enfield, N. J., & Sidnell, J. (2022). *Consequences of language: From primary to enhanced intersubjectivity*. The MIT Press.

Ferrara, L., & Hodge, G. (2018). Language as description, indication, and depiction. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00716

Feyaerts, K., Rominger, C., Lackner, H. K., Brône, G., Jehoul, A., Oben, B., & Papousek, I. (2022). In your face? Exploring multimodal response patterns involving facial responses to verbal and gestural stance-taking expressions. *Journal of Pragmatics*, *190*, 6–17. https://doi.org/10.1016/j.pragma.2022.01.002

Gabarró-López, S. (2024). Towards a description of PALM-UP in bidirectional signed language interpreting. *Lingua*, *300*. https://doi.org/10.1016/j.lingua.2023.103646

Goffman, E. (1963). *Behavior in Public Places*. Free Press. http://ebookcentral.proquest.com/lib/hioa/detail.action?docID=4934774

Goffman, E. (1981). *Forms of Talk*. University of Pennsylvania Press.

Goffman, E. (1986). *Frame analysis: An essay on the organization of experience*. Northeastern University Press.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. Academic Press.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, *32*(10), 1489–1522. https://doi.org/10.1016/S0378-2166(99)00096-X

Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse & Society*, *18*(1), 53–73. https://doi.org/10.1177/0957926507069457

Goodwin, M. H. (1980). Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry*, *50*(3–4), 303–317. https://doi.org/10.1111/j.1475-682X.1980.tb00024.x

Hansen, J. P. B., & Svennevig, J. (2021). Creating space for interpreting within extended turns at talk. *Journal of Pragmatics*, *182*, 144–162.

Haug, T., Bontempo, K., Leeson, L., Napier, J., Nicodemus, B., Van Den Bogaerde, B., & Vermeerbergen, M. (2017). Deaf leaders' strategies for working with signed language interpreters: An examination across seven countries. *Across Languages and Cultures*, *18*(1), 107–131. https://doi.org/10.1556/084.2017.18.1.5

Heath, C. (1986). *Body Movement and Speech in Medical Interaction*. Cambridge University Press. https://doi.org/10.1017/CBO9780511628221

Henley, R., & McKee, R. (2020). Going through the motions: Participation in interpreter-mediated meeting interaction under a deaf and a hearing chairperson. *International Journal of Interpreter Education*, *12*(1).

Janzen, T., & Shaffer, B. (2023). *Signed Language and Gesture Research in Cognitive Linguistics*. Walter de Gruyter.

Janzen, T., Shaffer, B., & Leeson, L. (2023). What I know is here; what I don´t know is somewhere else: Deixis and gesture spaces in American Sign Language and Irish Sign Language. In T. Janzen & B. Shaffer (Eds.), *Signed language and gesture research in cognitive linguistics.* (pp. 211–242). De Gruyter.

Johnston, T. (2019). *Auslan Corpus Annotation Guidelines*. Macquarie University.

Jucker, A. H., Hausendorf, H., Dürscheid, C., Frick, K., Hottiger, C., Kesselheim, W., Linke, A., Meyer, N., & Steger, A. (2018). Doing space in face-to-face interaction and on interactive multimodal platforms. *Journal of Pragmatics*, *134*, 85–101. https://doi.org/10.1016/j.pragma.2018.07.001

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Kendon, A. (1990). *Conducting Interaction. Patterns of behavior in focused encounters.* Cambridge University Press.

Lang, R. (1978). Behavioral Aspects of Liaison Interpreters in Papua New Guinea: Some Preliminary Observations. In *Language Interpretation and Communication* (pp. 231–244). Springer US. https://doi.org/10.1007/978-1-4615-9077-4_21

Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society*, *32*(2), 177–201. https://doi.org/10.1017/S004740450332202X

Liddell, S. K. (1980). *American sign language syntax* (Reprint 2021., Vol. 52). Mouton Publishers.

Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press.

Llewellyn-Jones, P., & Lee, R. G. (2014). *Redefining the role of the community interpreter: The concept of role-space*. SLI Press.

Mason, I. (2012). Gaze, positioning and identity in interpreter-mediated dialogues. In C. Baraldi & L. Gavioli (Eds.), *Coordinating participation in dialogue interpreting* (pp. 177–199). John Benjamins Publishing.

Mondada, L. (2006). Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis. In H. Knoblauch, J. Raab, H.-G. Soeffner, & B. Schnettler (Eds.), *Video Analysis* (pp. 1–18). Lang.

Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, *9*(2), 194–225. https://doi.org/10.1177/1461445607075346

Mondada, L. (2013). *Interactional space and the study of embodied talk-in-interaction* (pp. 247–275). De Gruyter. https://doi.org/10.1515/9783110312027.247

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, *65*, 137–156. https://doi.org/10.1016/j.pragma.2014.04.004

Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, *51*(1), 85–106. https://doi.org/10.1080/08351813.2018.1413878

Napier, J. (2007). Cooperation in interpreter-mediated monologic talk. *Discourse & Communication*, *1*(4), 407–432. https://doi.org/10.1177/1750481307082206

Napier, J., Carmichael, A., & Wiltshire, A. (2008). Look-pause-nod: A linguistic case study of a deaf professional and interpreters working together. In *Deaf Professionals and Designated Interpreters: A New Paradigm* (pp. 22–42). Gallaudet University Press.

Napier, J., Oram, R., Young, A., & Skinner, R. (2019). When I speak people look at me. *Translation and Translanguaging in Multilingual Contexts*, *5*(2), 95–120. https://doi.org/10.1075/ttmc.00027.nap

Nilsson, A.-L. (2010). *Studies in Swedish Sign Language: Reference, Real Space Blending, and Interpretation* [Stockholms Universitet, Humaistiska fakulteten, Institutionen för lingvistik]. https://su.diva-portal.org/smash/record.jsf?pid=diva2%3A305811&dswid=-6372

Nilsson, A.-L. (2023). Exploring real space blends as indicators of discourse complexity in Swedish sign language. In T. Janzen & B. Shaffer (Eds.), *Signed Language and Gesture Research in Cognitive Linguistics* (pp. 275–301). De Gruyter.

Pasquandrea, S. (2011). Managing multiple actions through multimodality: Doctors' involvement in interpreter-mediated interactions. *Language in Society*, *40*(4), 455–481. https://doi.org/10.1017/S0047404511000479

Poignant, E. (2021). The cross-lingual shaping of narrative landscapes: Involvement in interpreted storytelling. *Perspectives, Studies in Translatology*, *29*(6), 814. https://doi.org/10.1080/0907676X.2020.1846571

Quinto-Pozos, D., Alley, E., Casanova de Canales, K., & Treviño, R. (2015). When a language is underspecified for particular linguistic features: Spanish-ASL-English interpreters' decisions in Mock VRS Calls. In *Signed language interpretation and translation research: Selected papers from the first international symposium* (pp. 212–234). Gallaudet University Press.

Rosenthal, A. (2009). Lost in transcription: The problematics of commensurability in academic representations of American Sign Language. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, *29*(5), 595–614. https://doi.org/10.1515/TEXT.2009.031

Rossano, F. (2010). Questioning and responding in Italian. *Journal of Pragmatics*, *42*(10), 2756–2771. https://doi.org/10.1016/j.pragma.2010.04.010

Rossano, F. (2012). Gaze in Conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (1st ed., pp. 308–329). John Wiley & Sons.

Roy, C. B. (2000). *Interpreting as a discourse process*. Oxford University Press.

Roy, C. B. (2011). *Discourse in Signed Languages*. Gallaudet University Press. http://ebookcentral.proquest.com/lib/hioa/detail.action?docID=4860681

Ruusuvuori, J., & Peräkylä, A. (2009). Facial and Verbal Expressions in Assessing Stories and Topics. *Research on Language and Social Interaction*, *42*(4), 377–394. https://doi.org/10.1080/08351810903296499

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Santiago, R. R., Barrick, L. F., & Jennings, R. (2015). Interpreters' views on idiom use in ASL-to-English Interpreting. In B. Nicodemus & K. Cagle (Eds.), *Signed language interpretation and translation research: Selected papers from the first international symposium* (pp. 164–195). Gallaudet University Press.

Skedsmo, K. (2021). *Troubleshooting in Norwegian Sign Language. A conversation analytic approach to other-initiations of repair in multiperson Norwegian Sign Language interaction, and ways to communicate conversation analytic data on signed languages in printed and online publications.* [PhD]. OsloMet - Oslo Metropolitan University.

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., & Rossano, F. (2010). A Scalar View of Response Relevance. *Research on Language and Social Interaction*, *43*(1), 49–56. https://doi.org/10.1080/08351810903471381

Sutton-Spence, R., & Woll, B. (1999). *The linguistics of British Sign Language*. Cambridge University Press.

Vranjes, J., & Bot, H. (2021). A multimodal analysis of turn-taking in interpreter-mediated psychotherapy. *The International Journal of Translation and Interpreting Research*, *13*(1), 101–117. https://doi.org/10.12807/TI.113201.2021.A06

Vranjes, J., & Brône, G. (2021). Interpreters as laminated speakers: Gaze and gesture as interpersonal deixis in consecutive dialogue interpreting. *Journal of Pragmatics*, *181*, 83–99. https://doi.org/10.1016/j.pragma.2021.05.008

Vranjes, J., Brône, G., & Feyaerts, K. (2018). On the role of gaze in the organization of turn-taking and sequence organization in interpreter-mediated dialogue. *Language and Dialogue*, *8*(3), 439–467. https://doi.org/10.1075/ld.00025.vra

Wadensjö, C. (1998). *Interpreting as interaction*. Longman.

Wadensjö, C. (2004). Dialogue interpreting: A monologising practice in a dialogically organised world. *Target: International Journal of Translation Studies*, *16*(1), 105–124. https://doi.org/10.1075/target.16.1.06wad

Wadensjö, C. (2017). Dialogue Interpreting and the Distribution of Responsibility. *Hermes (Århus, Denmark)*, *8*(14), 111. https://doi.org/10.7146/hjlcb.v8i14.25098

Young, L., Morris, C., & Langdon, C. (2012). "He Said What?!" Constructed dialogue in various interface modes. *Sign Language Studies*, *12*(3), 398–413.

9. Appendix

Main tiers	Transcript conventions	Explanation
GLOSS		Identifies tokens of lexical signs that are part of NTS source utterances.
INT Norwegian		An orthographic transcription of the interpreter's verbal rendition into Norwegian
Trans		A translation into English. Source utterances and renditions are both provided with an English translation.
	<DM:example>	Identifies a discourse marker
	POSS-1P.s	Identifies first person singular possessive pronoun
	PRO-1P.s	Identifies first person singular personal pronoun
	PRO-2P.s	Identifies second person singular personal pronoun
	Raised eyebrows------	Descriptions of embodied actions are delimited between * * Transcriptions of embodied actions are based on Mondada (2018)
Gaze	*Cora----*	Identifies gaze towards named interlocutor for as long as dashes show *indicates the point where gaze shifts
	ss---------	Identifies gaze towards signing space for as long as dashes show
	[[	Identifies points of simultaneity between source utterance and rendition
	*---->	Action described continues across subsequent lines

	*----->>	Action described continues until and after extract ends
	#	Indicates the exact moment at which the screen shot has been recorded
	(.)	Identifies pause lasting less than 0.3 seconds
Head G		Identifies a head gesture
	side-turn/downward--	Identifies types of head gestures. Transcriptions of head gestures are based on Allwood et al. (2007)

Vibeke Bø
OsloMet – Oslo Metropolitan University
Pilestredet 42
NO-0130 Oslo
Norway

vibeb@oslomet.no

Biography: Vibeke Bø is an interpreter, interpreter trainer and interpreting researcher with experience in linguistics, interaction studies and interpreting studies. She holds a Master's degree in linguistics from 2010, where she investigated a syntactic phenomenon in Norwegian Sign Language (NTS). Recently, she submitted her PhD dissertation, focusing on the semiotic practices of signed-to-spoken interpreting. This dissertation exemplifies her interdisciplinary approach. From this work, three articles are in the process of being published in peer-reviewed journals. She currently teaches in the BA program in Norwegian Sign Language at OsloMet.

Parallèles 37(1) April 2025

Reformulation structures in French Belgian Sign Language (LSFB) > French interpreting: A pilot multimodal study

Sílvia Gabarró-López

Universitat Pompeu Fabra & Université de Namur

Abstract

This article examines reformulation structures when interpreting French Belgian Sign Language (LSFB) into spoken French. Reformulation structures are defined as two segments of discourse, where the first segment conveys a message and the second segment, introduced by a marker, expresses the same message differently. By adopting a multimodal approach, interlingual reformulation structures (between the source and the target languages) and intralingual reformulation structures (within the target language) are described, focusing on their distribution, form, and semiotic composition. The dataset comprises dialogues produced by two LSFB signers and their interpretations into French by two LSFB–French interpreters. Interlingual and intralingual reformulation structures are present in French interpretations, although less frequently than in LSFB source dialogues. The most frequent forms of reformulation structures are found in both datasets. Interpreters do not seem to be influenced in their gesturing by the signs produced in the source LSFB dialogues. Still, they engage their hands, bodies, faces, and voices in their renditions. Hence, interpreters draw on all their available semiotic resources to convey meaning but differ from how source signers do it. In future research, the dataset should be enlarged and the type of manual gestures and nonmanual articulators used should be more closely investigated.

Keywords

Reformulation structures, semiotic repertoires, interpreting, French Belgian Sign Language (LSFB), French

 DOI:10.17462/para.2025.01.10

1. Introduction

Reformulation is pervasive in spoken and signed languages (e.g., Blakemore, 1993; Cuenca & Bach, 2007; Cuxac, 2007; Meurant, 2022). This phenomenon has been attested in prepared and unprepared discourses, in monological, and conversational settings. That is, reformulation is intrinsic to all types of human communication, whether written, spoken, or signed. Although most people would agree that reformulation involves saying the same thing differently, the concept of reformulation may have different implications depending on the field of study.

On the one hand, reformulation has traditionally been defined in the field of linguistics as the process wherein two segments of discourse (X and Y) provide the same information using different words/signs or expressions in a language. This semantic equivalence established between X and Y is called paraphrastic reformulation and is illustrated in example (1)[1], taken from Meurant et al. (2022, p. 324). The speaker has been asked to describe a picture. She explains that because of an optical illusion, there are two possible perspectives from which the picture can be looked at and interpreted.

(1) *Ça se joue sur euh l'illusion optique, c'est-à-dire que euh il y a deux perspectives.*
<X1> M1 <Y1>
< It plays on uhm optical illusion, > <u>that is to say</u> < uhm there are two perspectives. >

A broader definition of the phenomenon under study is nonparaphrastic reformulation, in which the Y segment is used to narrow, expand, adjust, specify, clarify, define, correct, or modify different aspects of the X segment (Murillo, 2016). Example (2), also taken from Meurant et al. (2022, p. 349), belongs to a conversation about what having a good level of French means. The speaker says that there is a difference between oral and written practices (X segment) and expands this statement by saying that you can have different levels in these two modalities (Y segment).

(2) *Déjà si tu considères l'oral ou l'écrit parce que*
<X1> M1
tu peux avoir un niveau de français qui est très différent selon que tu le pratiques à l'oral ou à l'écrit donc.
<Y1>
< If you consider the oral or the written > <u>because</u> < you can have a very different level of French depending on whether you practice it orally or in writing so. >

Regardless of the types of reformulation structures, they can be marked and unmarked. In unmarked reformulation structures, there is no reformulation marker. That is to say, the two segments are not connected by a word or combination of words functioning as markers. Marked reformulation structures may have different types of markers. Traditionally, these markers have been characterized as either introducing paraphrastic reformulation (e.g., *c'est à dire que* 'that is to say' in example (1)) or nonparaphrastic reformulation (e.g., *in fact*). Nevertheless, the polyfunctional nature of markers is such that those that have traditionally been classified as paraphrastic are found in nonparaphrastic structures or that the marker of a reformulation structure does not typically belong to the realm of reformulation, as *parce que* 'because' in example (2) (Pons Bordería, 2013).

[1] French examples are written in italics. Below each line, the form of reformulation structures is presented. The translations into English are provided below (the different segments of the reformulation structure are delimited with angled brackets and the marker is underlined).

On the other hand, reformulation may be understood as the mechanism used by translators and interpreters to bridge the communicative divide between languages and their respective cultures. In this context, reformulation is found in any translation or interpretation, as it involves conveying the meaning of a text/discourse in the source language using the words/signs and the structures of the target language. This type of reformulation is what Jakobson (1963) calls 'interlingual reformulation', which involves modifications in syntax, semantics, and pragmatics across languages. However, translations and interpretations also have instances of 'intralingual reformulation' (Jakobson, 1963), that is, cases in which reformulation takes place within the target language for different reasons, such as the lack of one-to-one correspondence for a concept between two languages.

In this paper, the concept of reformulation includes paraphrastic and nonparaphrastic structures as well as interlingual and intralingual reformulation structures. In what follows, previous research on the type of reformulation structures produced by interpreters is presented alongside a theoretical framework that supports the analysis of different human communicative practices.

1.1. Interlingual and intralingual reformulations, description and depiction

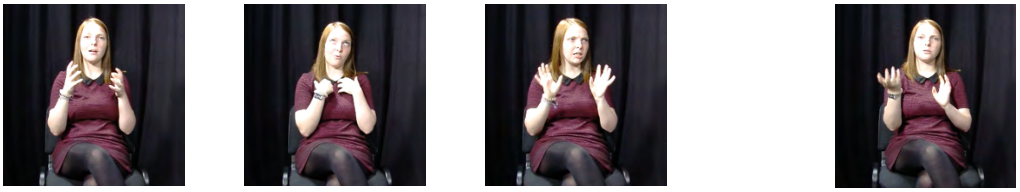
Translated and interpreted texts/discourses are composed of different types of reformulations, which may offer insights into the cognitive processes of translators and interpreters and the strategies they employ to convey meaning accurately and effectively. However, reformulation has been scarcely studied in the field of translation and interpreting. Using voice-recorded data, Woroch (2010) describes paraphrastic and nonparaphrastic reformulation structures produced by interpreters who work from French to Polish. She first examines reformulations in source French texts and then reformulations in target Polish texts. By comparing the source and target productions, she teases apart interlingual reformulations from intralingual reformulations. After her analysis, Woroch (2010) concludes that interlingual and intralingual reformulations add value to interpreted renditions, making the target Polish discourse more accessible to the audience.

Woroch (2010) provides a comprehensive account of the different types of reformulation structures that she found in conference interpreting, so her research can be a good starting point for a replication study using another pair of languages. However, if we understand language as multimodal, we need another theoretical framework with which the other semiotic resources available to speakers/signers and interpreters can be accounted for, including the manual and nonmanual activity. Following Peirce (1955) and Clark (1996), Ferrara & Hodge (2018) propose that spoken and signed communication involves three modes of signaling:

- *Description* includes "lexicalized manual signs of deaf signed languages [see examples in Figure 5, from pictures 2 to 8], the spoken or written words of spoken languages [see examples (1) and (2)], culturally-specific emblematic manual gestures such as the OK and THUMBS-UP gestures […], and conventionalized intonation contours [e.g., the intonation of a question]" (Ferrara & Hodge, 2018, p. 3).
- *Indication* is defined as indexing referents with a variety of "forms such as the English function words *it* and *this*, as well as hand-pointing, lip-pointing, and other culturally-specific bodily actions during which speakers or signers extend parts of their body (or objects that act as an extension of their body) in a direction toward, or contacting, some referent in the context of the utterances" (Ferrara & Hodge, 2018, p. 4).
- *Depiction* may include tokens with "[varying degrees] of conventionalization across a community" (Ferrara & Hodge, 2018, p. 5), such as depicting signs in signed languages or metaphoric manual gestures in spoken languages, as well as the enactment of the

actions, words or thoughts of a referent (which could be prior acts of description or indication).

Hence, these three modes of signaling are not exclusive categories. For instance, Figure 1 illustrates an example in which the speaker criticizes the point of view of the Académie française and combines description and depiction in the two segments of a reformulation structure for this purpose. Most of the time, the mode of signaling is description. However, when she says '*la langue c'est sacré'* ('language is sacred'), *'féminiser c'est complètement absurde'* ('feminizing is completely absurd'), and '*on va tuer la langue française'* ('we are going to kill the French language'), she enacts the point of view of the Académie française in a dramatic way using her intonation together with movements of both hands, the head and the chest, and her facial expression.



Et il faut pas euh je pense dire que la langue c'est sacré que, par exemple,

<X1> M1

avec la féminisation euh des noms de métier, des titres et tout ça, euh quand on entend le point de vue de l'Académie française, c'est quand même un peu aberrant où il en ils en viennent à dire que

<Y1

féminiser c'est complètement absurde. « On va tuer la langue française ! ».

Y1>

< And one mustn't uh I think say that language is sacred that, > <u>for instance</u>, < with the feminization uhm of job titles, diplomas and all that, uhm when you hear the point of view of the Académie française, it's a bit aberrant when he they say that feminizing it's completely absurd. "We'll kill the French language!" >

Figure 1. Excerpt of a dialogue in French in which the speaker combines description and depiction (adapted from Meurant et al., 2022, pp. 349–350)

The choice of this theoretical framework (Ferrara & Hodge, 2018) for the present research is motivated by the fact that not only does it allow the comparison of spoken and signed languages, but it has also been used for the study of different phenomena in sign language interpreted renditions and in both interpreting directions (e.g., Meurant et al., 2022; Bø, in press; and Bø, this volume).

1.2. Objectives and hypotheses

To the best of my knowledge, reformulation structures have not been studied using multimodal interpreted spoken data and have been scarcely studied in sign language interpreting (Meurant et al., 2022). This paper addresses these shortcomings by describing interlingual and intralingual reformulation structures in signed-to-spoken language interpreting using multimodal data, i.e., video recordings. Furthermore, this paper will add to the small body of research on the signed-to-spoken language direction in interpreting, which has received less attention than the spoken-to-signed language direction so far (Wang, 2021). The languages under study are French Belgian Sign Language (LSFB) and spoken French (Belgian variety). Both languages are used in Wallonia (southern Belgian region) and in Brussels (where they coexist with Flemish Sign Language (VGT) and spoken Flemish). In these two regions, LSFB remains a minority and minoritized language.

The objectives of this paper are threefold:

1. To study the distribution of reformulation structures. In line with Woroch (2010), this phenomenon is first identified in source LSFB data and later in target French data so that interlingual and intralingual reformulation structures can be teased apart.
2. To describe the form of reformulation structures. Once identified, the arrangement of the X and Y segments, the position of the markers, and their form (e.g., lexicalized signs, pause fillers, etc.) are detailed.
3. To examine the semiotic composition of reformulation structures and the modes of signaling. In other words, the interplay between the manual and nonmanual activity in LSFB and the interplay between speech and the manual and nonmanual activity (i.e., eye gaze direction, facial expressions, and head and body movements) in French to describe and depict[2].

Two hypotheses are formulated. The first one is that interpreters may use fewer reformulation structures in their productions than LSFB signers, given the cognitive demands placed on interpreters such as memory, cognitive load, and time lag. The second hypothesis is that interpreters will incorporate the signs and gestures of source signers in their reformulations in line with Janzen et al. (2016).

In the remainder of this paper, the dataset used for this research is presented along with how reformulation structures were identified and characterized, and how videos were annotated. Afterward, the distribution of reformulation structures, their form, and their semiotic composition for source LSFB and target French data are described and later compared. Finally, the relationship between reformulation structures in target discourse and interpreting strategies is also discussed, as well as the implications of this paper for the field of interpreting research and training.

2. Methodology

2.1. The two datasets

This study draws on corpus data that were recorded in a studio setting. The source data were extracted from the LSFB Corpus (Meurant, 2015), namely the reference corpus for this sign language. It includes 100 signers from different places in Belgium where LSFB is used. There is a balance among signers regarding gender, age, and linguistic background. Participants not only provided dialogical signed data, but they were also asked to fill in a metadata form. Before the recordings, they also signed an informed consent allowing the recorded data to be made openly

[2] For this pilot study, indication is not analyzed in line with Meurant et al. (2022).

available on the corpus website, but not their metadata (which is restricted to researchers). The LSFB Corpus data are very convenient because deaf annotators previously annotated the signs produced in the dialogues and professional translators translated the videos into written French, in case future contrastive research between translated and interpreted data were to be conducted. Two dialogues between two female deaf LSFB signers (S055 and S056) who recount childhood memories and discuss issues related to the differences between deaf and hearing cultures were used for this paper. This small dataset of source LSFB dialogues totals 10 minutes (see Table 1).

Topic of discourse	Data	Participants	Duration
Childhood memories	Source LSFB	S055 & S056	4’53”
	Target French	I002	4’56”
		I006	4’51”
Cultural issues	Source LSFB	S055 & S056	4’46”
	Target French	I002	5’02”
		I006	4’58”

Table 1. Description of the dataset

The target data were taken from the CorMILS Pilot Project (Gabarró-López, 2018), which contains interpreted data by the first cohort of final-year students of the Master’s degree in LSFB – French interpreting of the UCLouvain. CorMILS’ data include the two interpreting directions, the two dialogues from the LSFB Corpus mentioned above, and two comparable dialogues from the FRAPé (Multimodal French) Corpus (Meurant et al., ongoing) used as source data. The six participants of the first cohort had different profiles, including two students with previous experience as interpreters in the educational setting and four non-experienced students. Similarly to the LSFB and FRAPé corpora, participants filled in a metadata form (which was inspired by the metadata forms used in these two corpora) and signed an informed consent form. Although they agreed to be recorded, their data are not openly available and can only be used by researchers.

The LSFB > French renditions produced by the two experienced students (I002 and I006) were selected for this pilot study, totaling 20 minutes. The two participants are a woman and a man, aged 30–40, with 5–6 years of interpreting experience in different educational institutions. This choice was motivated by the recording conditions. Participants were shown the videos twice. The first time they watched the source data and could ask questions about the meaning of signs or the signs to be used for a particular expression. Afterward, participants were shown the videos a second time and had to interpret them. As could be expected, non-experienced participants struggled while interpreting because of the speed of natural dialogues which could not be stopped, so they produced more omissions or interpreting errors than experienced participants. Therefore, it was decided to keep the data of the two experienced participants to avoid bias. Despite the small size of the dataset, which does not allow broader generalizations to be made, the foundations for the description of reformulation structures can be laid so that future research may build on them.

2.2. Annotation procedure: Identifying reformulation structures and characterizing them

After closely inspecting the videos of source and target data, they were annotated with ELAN (Lausberg & Sloetjes, 2009) following a three-step process: reformulation structures

were identified, their content was transcribed/summarized[3], and the articulators used for description and depiction were annotated. Once all reformulation structures were annotated in the source and target discourses, they were classified in the target renditions as interlingual (if they had been produced in LSFB) or intralingual (if they were only uttered in French).

Although reformulation structures have extensively been described in the literature, their identification in the wild is not a straightforward process, as there are other neighboring phenomena such as elaboration or false starts which are difficult to tease apart. Hence, a clear set of criteria, such as those proposed by Meurant et al. (2022: 329), was needed:

> On the one hand, it implies that, between the source and the reformulated statement there is something identical and something different. This makes it possible to distinguish reformulation from repetition (Tannen, 1989). On the other hand, since reformulation is based on the creation of an equivalence between two utterances, the act of reformulation implies a reflexive, or metalinguistic return to the first statement. This makes it possible to distinguish reformulation from all cases where the sequence of statements, from one to the next, advances the information, maintaining a common core to which new information is added.

Furthermore, only reformulation structures introduced by a marker were analyzed to ensure the comparison between the source and target data and the replicability of the present study. Regarding markers, they were identified on the go. When a chunk of discourse matched the definition of reformulation, the marker (if any) was identified. No *a priori* distinctions were made between the markers, that is, they could be produced by any articulator and be found in any position (Meurant et al., 2022). Before starting the annotation of the whole dataset, reformulation structures were first annotated by the author[4] and checked by another researcher for the first two minutes of I002 renditions to enhance the application of the three criteria mentioned above and to sort out ambiguities.

2.3. Annotation template

The manual activity of LSFB dialogues had previously been annotated using ID-glosses[5] and been translated into written French (see 2.1). However, interpreted French data had not received any annotation, neither for speech nor for manual and nonmanual behavior. To ensure comparability, a common annotation template was created in ELAN for source and target data, including four tiers (each of which was preceded by the signer or interpreter's code):

- *Refor_XY*: this tier was used to capture the scope of the reformulation structure, i.e., where the X and Y segments and the marker (M) of each reformulation structure started and ended. In the annotation, the X, the Y, and the M of a given reformulation structure were followed by the same number as in <X1> M1 <Y1>. The brackets surrounding the letters allow us to visualize where the marker was placed, as in <X1> <Y1 M1 Y1> (meaning that the marker is embedded in the Y segment) or whether the Y segment of a reformulation structure was the X segment of the next one, as in <X1> M1 <Y1 X2> M2 <Y2>.

[3] Although source LSFB data were previously annotated using glosses, the main ideas stated in the reformulation structures were summarized in written French to grasp easily what they were about. By contrast, target French data were not previously annotated, so what was said in the reformulation structure was fully transcribed.

[4] I am a hearing fluent user of LSFB and French. I am an academic trained first in translation and interpreting and later in linguistics, but I have not worked as a sign language interpreter.

[5] ID-glosses consist of words written in capital letters. They are used to label a sign consistently regardless of the context in which it appears (Johnston, 2010).

- *Refor_Content*: this tier contained a written summary or transcription of what had been signed or spoken for each segment of the reformulation structure and its marker. For the latter, additional information could be added in the annotation if, for example, the marker was spoken, and a gesture was produced simultaneously.
- *Refor_Description*: this tier comprises the annotation of the articulators, one after another, that signal description.
- *Refor_Depiction*: this tier comprises the annotation of the articulators, one after another, that signal depiction.

The articulators were annotated using the different abbreviations presented in Table 2.

Abbreviations	Articulators
MD	Movement of the right hand
MG	Movement of the left hand
VX	Use of speech
TE	Head movement
EX	Facial expression
BU	Body movement
RE	Eye gaze direction
MO	Mouth gesture
LA	Mouthing

Table 2. Abbreviations used to annotate the articulators (Meurant et al., 2022, p. 331)

The files containing target French renditions had an additional tier called *Refor_type* (also preceded by the interpreters' codes) in which it was annotated whether the reformulation structure was interlingual or intralingual (see Figure 2).

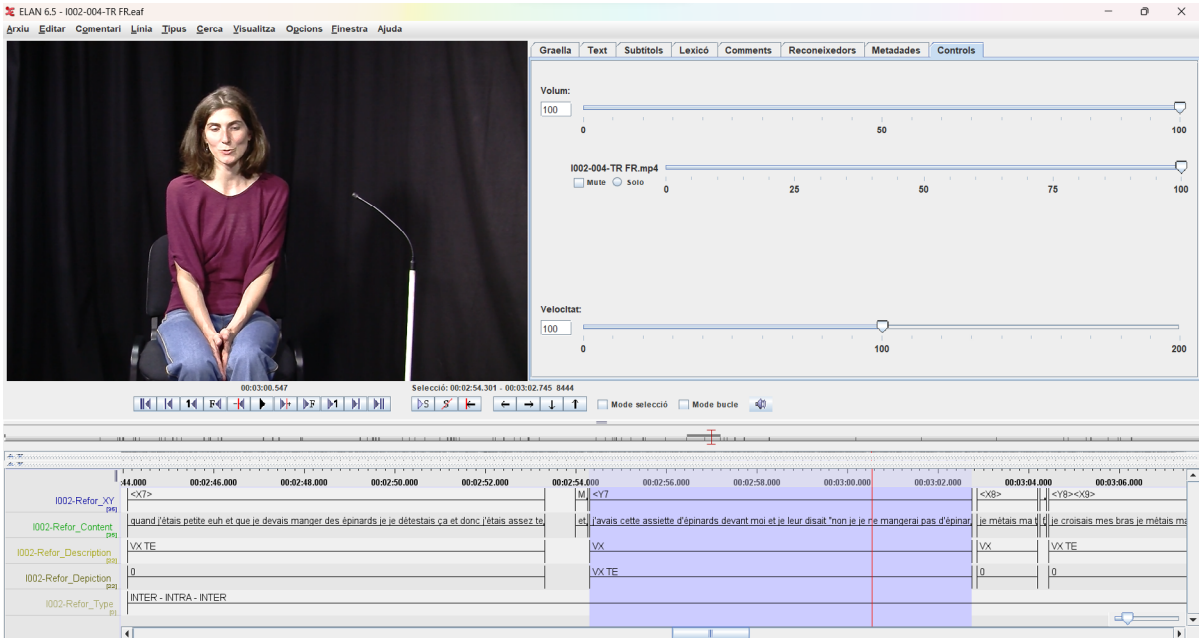


Figure 2. Screenshot of a target French data file

Once source and target videos were completely annotated, the annotations were extracted using Excel files for analysis.

3. Results

This section is divided into two parts: 3.1 is devoted to source LSFB data and 3.2 to target French data. Both parts have the same structure, including the number of reformulation structures produced by signers or interpreters, the form of these structures, the types of explicit markers used, the modes of signaling employed, and the semiotic composition of the X and Y segments.

3.1. Reformulation in source LSFB data

There are 25 reformulation structures in the source LSFB dataset: 15 were produced by S055 and 10 by S056. These reformulation structures can be independent, combined, or embedded with other reformulation structures depending on the arrangement of segments, as shown in Table 3.

Type of reformulation structures	Arrangement of segments	Number of cases
Independent	<X1> M1 <Y1>	14
Combined	<X1> M1 <Y1 X2> M2 <Y2>	5 (grouped in two reformulation structures of 3 and 2, respectively)
Embedded	<X1> M1 <Y1<X2> M2 <Y2> Y1>	6 (grouped in two reformulation structures of 3, respectively)

Table 3. Number of reformulation structures per type in source LSFB data

Most reformulation structures are independent (14 cases), and the marker is placed between the X and the Y as in example (3)[6], in which S055 recounts a memory of her friends at school.

(3) FRIEND HEARING YES SCHOOL PT:LOC HAVE SEVERAL
<X1> M1 <Y1>
< I also had hearing friends, > yes, < at school I had several hearing friends. >
CLSFBI2703_S055_02:58.495 – 03:00.973

The marker is placed after the Y segment only in one reformulation structure, represented in example (4). S056 mentions one of the differences between deaf and hearing cultures, namely the differences when calling a deaf or a hearing person.

[6] LSFB examples are written in small capital letters, as established in the sign language literature. For long examples, omissions are marked with […]. PT stands for a pointing. If PT is followed by two points and another word, the latter defines whether the pointing is used as a first-person singular pronoun (PRO1) or to indicate a location (LOC). GSIGN is used for manual forms that are not lexicalized signs, i.e., gestures. Glosses composed of two words are separated by a hyphen (e.g., DEAF-CLUB). Below the glosses of each example, the form of the reformulation structure, the translation into English (following the same conventions of the examples in French, cf. footnote 1), and the reference of the example (session, task, signer, and time code) are presented. Bear in mind that the translation of some markers such as PALM-UP or PT are tentative as they do not have a one-to-one correspondence in spoken English or French.

(4) UNDERSTAND NOT CULTURE DIFFERENT HEARING DEAF DIFFERENT PLACE WAY CALL PALM-UP

<X1> <Y1> M1

< They don't understand that that's a different culture, > <deaf and hearing people call their peers differently, > indeed.

CLSFBI2704_S056_01:43.504 – 01:46.920

The remaining 11 reformulation structures do not stand alone. On the one hand, five reformulation structures are instances of a combination of the Y segment of the first reformulation with the X segment of the following one (i.e., <X1> M1 <Y1 X2> M2 <Y2>). These five cases are combined in two structures; one has two reformulation structures, and the other has three. In example (5), S055 explains why eye gaze is important for deaf children and their parents.

(5) ALSO LOOK ALSO PT […] PARENTS HEARING KNOW-NOT HOW CHILD ALSO LOOK PALM-UP

<X1> M1 <Y1 X2> M2

[…] PERSON DEAF KNOW HOW OR DEAF TEACH ON HEARING DO

<Y2>

< The eye contact is also important, > I mean, < […] hearing parents don't know how to make their kids look at them. > In fact, < […] a deaf person knows how to do it, or a deaf person should teach hearing parents how to do it. >

CLSFBI2704_S055_01:47.444 – 02:07.048

On the other hand, six reformulation structures are instances of embedment. In other words, the Y segment of the first reformulation has two reformulation structures embedded, similarly to <X1> M1 <Y1<X2> M2 <Y2> <X3> M3 <Y3>Y1>. Most embedded reformulation structures have the marker between the X and the Y segment, except for one reformulation structure in which the marker is placed after the Y segment. These two possibilities are shown in example (6), in which S056 states her preference for the deaf club over the cinema.

(6) CULTURE ALSO FOR PT:PRO1 ALSO GO CINEMA OR GO DEAF-CLUB DEAF DIFFERENT PALM-UP

<X1> M1

PERSON HEARING FEELING LOVE CINEMA DEAF […] NOT NEED BECAUSE […] BEFORE LITTLE NOT SUBTITLES

<Y1 <X2>

PT:PRO1 NOT NEED GO LITTLE UNDERSTAND NOT PICTURE ALSO LEAVE PT

<Y2> M2

MORE DEAF-CLUB GOOD BECAUSE DEAF SIGN-LANGUAGE THERE GSIGN MORE COMMUNICATION ALSO

<X3> M3 <Y3> Y1>

< Another cultural difference is what to choose between going to the cinema or the deaf club. > So < hearing people love going to the cinema, but deaf people don't [because it's stupid]. When I was a child, I didn't want to go to the cinema because there weren't subtitles. > < It's not worth going there to see pictures without understanding, > so < going

to the deaf club was a better option because I could sign, > I mean, < there was more communication. >

CLSFBI2704_S056_03:44.730 – 04:05.870

As shown above, different manual forms are used as markers. The most frequently used markers are gestural forms such as the PALM-UP gesture (Figure 3) and GSIGN (i.e., wiggling or rubbing fingers, which are used as pause fillers in LSFB), and partly-lexicalized signs, i.e. pointings (Figure 4). Different lexicalized signs may also be used as reformulation markers and combinations of forms (see Table 4).

Figure 3. PALM-UP gesture

Figure 4. PT (pointing)

Type of marker	Form	Number
Lexicalized signs	YES	2
	BECAUSE	2
	IT-MEANS	2
	EXAMPLE	2
	ALSO	1
	BUT	1
Partly-lexicalized signs	PT	3
Gestures	PALM-UP	6
	GSIGN	4
Combinations	GSIGN PT GSIGN PALM-UP	1
	WHY BECAUSE	1

Table 4. Type, form, and number of markers used in source LSFB dialogues

Regarding the modes of signaling, description is present in all 25 LSFB reformulation structures. The semiotic resources recruited for description within these utterances include the two hands, head movements, and mouthings, whereas facial expressions (e.g., raising eyebrows to ask a question), body movements (e.g., body tilts to present two alternatives) and changes in eye gaze direction (e.g., to place referents in the signing space) are frequent but do not appear in all cases. In 12 reformulation structures, description is combined with depiction, either in the X or Y segments. The articulators used within these utterances for depiction include the two hands, facial expressions, and head and body movements, while changes in eye gaze direction, and mouth gestures are sometimes employed for this purpose, but not in a systematic way. Figure 5 illustrates an excerpt from example (6), particularly the third line of ID-glosses, in which description and depiction are combined.

'When I was a child, I didn't want to go to the cinema because there weren't subtitles. It's not worth going there to see pictures without understanding.'

CLSFBI2704_S056_03:44.730 – 04:05.870

Figure 5. Combination of description and depiction in the same reformulation segment

In this figure, S056 explains her experience as a child in the cinema. She uses description in the first four and the last two pictures. To this end, she articulates signs with her hands, moves her head, and produces mouthings. However, from the fifth to the eighth pictures, she depicts herself when she was a child in the cinema. In addition to the articulators used for description, S056 uses facial expressions, moves her body, and changes her eye gaze direction.

3.2. Reformulation structures in target French data

In target French discourse, I002 produces 16 reformulation structures (12 interlingual and four intralingual) and I006 produces 11 (six interlingual and five intralingual). Similarly to what is found in source LSFB data (see 3.1), reformulation structures can be independent, combined, or embedded (see Table 5).

Type of reformulation structures	Arrangement of segments	Number of cases
Independent	<X1> M1 <Y1>	16
Combined	<X1> M1 <Y1 X2> M2 <Y2>	2 (which made one reformulation structure)
Embedded	<X1> M1 <Y1<X2> M2 <Y2> Y1>	9 (grouped in three reformulation structures of 3)

Table 5. Number of reformulation structures per type in target French data

The most frequent form of reformulation structures in target French renditions is also <X1> M1 <Y1> (12 occurrences), as shown in example (7). I002 interprets the excerpt presented in example (5), producing an interlingual reformulation structure.

(7) *Je veux dire aussi qu'en termes de regard euh il y a aussi quelque chose de différent.*

<X1>

Par exemple,

M1

un professeur, des parents sourds (savent) à quel point le regard est important, qu'il faut apprendre à ce que les enfants puissent fixer le regard, alors qu'un entendant ne sait pas spécialement, il sait pas comment faire.

<Y1>

< I want to add that in terms of eye contact erm there is something different. > For instance, < a teacher, deaf parents know to what extent eye contact is important, and that children need to be taught to keep eye contact, whereas a hearing person doesn't necessarily know how to do it. >

CorMILS_I002-004-TR FR_02:03.194 – 02:19.916

Independent reformulation structures can have the marker embedded in the Y segment (four occurrences), as in example (8). In this excerpt, which follows the previous one, there is a change of speaker (i.e., X1 corresponds to S055 and Y1 to S056). When I002 interprets it, she produces an intralingual reformulation structure not reproduced from the source LSFB discourse.

(8) *Et c'est vrai que si si euh s'il n'y a pas ce lien avec le regard, ça peut devenir très violent pour l'enfant.*

<X1>

Oui, c'est vrai, la la communication, en fait,

<Y1 M1

n'y est pas et oui, c'est une forme de violence.

Y1>

< And it's true that if if erm if there is not this eye contact, it can become very violent for the kid. > < Yes, it's true, the the communication, in fact, is not there and yes, it's somehow violent. >

CorMILS_I002-004-TR FR_02:27.860 – 02:38.817

There is only one combined reformulation composed of two reformulation structures in target French discourse, illustrated in example (9). It also has the <X1> M1 <Y1 X2> M2 <Y2> form, as in example (5) of source LSFB data. In (9), I006 interprets S056's experience with the Scouts, producing two chained intralingual reformulation structures.

(9) *Je faisais aussi partie des scouts autant que lui. Emmm…*

<X1> M1

C'était le temps de partir en camp scout, et je faisais partie d'une troupe, […] euh

<Y1 X2> M2

je suis partie en camp pour voir un peu comment ça se passe la vie de scout.

<Y2>

< I took part in the Scouts as he did. > Mmm… < It was time to go camping with the Scouts, and I was part of a group, […] > erm < I went camping to see a little bit how life with the Scouts is. >

CorMILS_I006-003-TR FR_00:18.363 – 00:41.273

In embedded reformulation structures, the embedding is found in the Y segment, which can either be preceded by the marker—as in example (10)—or have the marker embedded after the last embedded reformulation—as in example (11). Each of these two examples puts together three reformulation structures. Example (10) is the only case where the marker appears after the first X segment. I002 interprets a memory of S055 related to lunchtime at home. She produces three reformulation structures: <X2> M2 <Y2> is intralingual, while the other two are interlingual.

(10) *Quand j'étais petite euh et que je devais manger des épinards, je je détestais ça, et donc j'étais assez têtue, mes parents aussi.*

<X1>

Et donc

M1

j'avais cette assiette d'épinards devant moi et je leur disais "non, je je ne mangerai pas d'épinards", et pour leur montrer euh ma détermination,

<Y1

je mettais ma tête sur mes mains, *fin,*

<X2> M2

je croisais mes bras, je mettais ma tête sur mes mains sur la table, *et ça veut dire que*

<Y2 X3> M3

je ne ne voyais pas ce qui se passait autour de moi, donc impossible de communiquer avec mon entourage […].

<Y3>Y1>

< When I was a kid erm and I had to eat spinach, I I hated it, and then I was quite obstinate, and so were my parents. > And then < I had this dish with spinach in front of me and I told them “no, I I won’t each spinach”, and to show my determination, < I put my head on my hands, > I mean, < I crossed my arms, I put my head on my hands on the table, > which means that < I couldn’t couldn’t see what was going on around me, so it was impossible to communicate with people around […]. >

CorMILS_I002-004-TR FR_02:44.089 – 03:21.033

Example (11) illustrates one of the two cases in which the marker of the main reformulation structure appears embedded in the Y segment. Interestingly, each case is produced by one interpreter and refers to the same moment in the source dialogue. In (11), I002 interprets how S055 used to perceive the deaf and hearing worlds. She produces three interlingual reformulation structures, as they are interpreted from the source dialogue.

(11) *Quand j'étais enfant, je me je me souviens de, en fait, très fort de deux mondes différents.*

<X1>

Il y avait le le monde sourd, euh

<Y1 <X2> M2

je quand j'accompagnais ma maman […], et puis mon papa […], donc j'avais euh fort un lien fort avec cette langue ainsi que ma sœur.

<Y2>

Mes parents tenaient tout de même à ce qu'on soit dans le monde entendant,

<X3>

et donc euh ils m'avaient inscrit à un cours de dessin […] avec les entendants.

M3 <Y3>

Eh donc J'étais vraiment partagée entre ces deux mondes.

M1 Y1>

< When I was a kid, I do I do remember, in fact, two very different worlds. > < There was the the deaf world, > erm < I when I went with my mum […], and then my dad [...], so I had erm a strong connection with this language as did my sister. > < My parents wanted anyway that we were in the hearing world, > and so erm < they sent me to drawing lessons […] with the hearing. > Erm so < I was in between these two worlds. >

CorMILS_I002-003-TR FR_02:09.500 – 03:02.940

In the target French dataset, three types of reformulation markers are found: connectors/ discourse markers, pause fillers, and combinations of connectors or connectors with pause fillers (see Table 6).

Type of marker		**Form**	**Number**
Connectors/discourse markers		*en fait* 'in fact'	4
		par exemple 'for example'	2
		oui 'yes'	2
		fin (enfin) 'well'	1
		et 'and'	1
		mais 'but'	1
Pause fillers		*euh* 'erm'	2
		emmm 'mmm'	1
Combinations	Connector + connector	*et puis en fait* 'and then in fact', *et donc* 'and so', *et voilà* 'and there you go', *et ça veut dire que* 'and it means that'	7
	Connector + pause filler	*et donc euh* 'and so erm', *euh donc* 'erm so', *et euh* 'and erm', *par exemple* emmm 'for example mmm'	6

Table 6. Type, form, and number of markers used in target French renditions

Regarding the modes of signaling, description is present in the 27 reformulation structures. The two interpreters use their voices to signal description in the utterances; that is, they produce words, sentences, and conventionalized intonation contours. Sometimes these tokens are combined with head movements (e.g., head tilts), facial expressions (e.g., eyebrows raised while asking a question or furrowed to express confusion), and hand gestures. In Kendon's (2004) terms, most of these gestures belong to the 'palm-up family' and only some to the 'palm-down family' (Figure 6). An example of the canonical form of a PALM-UP gesture is shown in Figure 3, but interpreters mostly produced one-handed or reduced forms (see Figures 7 and 8). These tokens are pragmatic gestures, which means that they relate to some aspects of discourse structure. For instance, the palm-down gesture's function is to "render unnecessary further action, inquiry or comment" (Kendon, 2004, p. 258) as expressed by I006 in example (12). While he produces the Y segment of the reformulation structure, he repeats the gesture three times (see underlined words).

(12) *Un jour où de nouveau je n'avais pas envie de manger, mes parents ont sans broncher m'ont dit de con... de manger mon assiette, sinon je ne me lèverais pas.*

'One day in which again I didn't want to eat, my parents told me without batting an eye to con… to eat my plate, otherwise, I would not leave the table.'

CorMILS_I006-004-TR FR_02:55.638 – 03:03.164

Figure 6. One-handed palm-down gesture

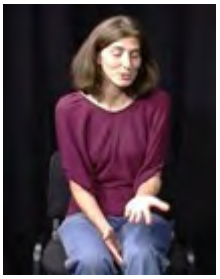


Figure 7. One-handed palm-up gesture

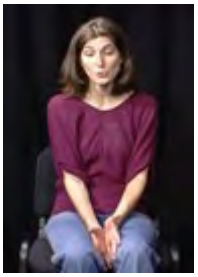


Figure 8. Reduced palm-up gesture

Reformulation segments in which depiction is combined with description are barely identified, only in one X segment and one Y segment of different reformulation structures. The hands are used in the former case to depict the word 'general', in which I006 draws a ball with his hands (see Figure 9). In the latter case, I002 uses her voice (i.e., she utters a sentence with a falling intonation contour) and head movements while she is enacting herself when she was a kid saying *'non, je je ne mangerai pas d'épinards'* ('no, I I won't eat spinach', see example (10) above).

Figure 9. Start and end position of the gesture depicting the meaning of 'general'

4. Discussion

The results presented in Section 3 show that fewer reformulation structures were used in target French discourse than in source LSFB discourse (16 and 11 reformulation structures produced by I002 and I006 respectively vs. 25 reformulation structures produced by S055 and S056 in the dialogues). These results indicate that the first hypothesis, namely a smaller number of reformulation structures in target French discourse due to interpreters' cognitive load and time lag (among other factors), is supported. However, it must be noted that only marked reformulation structures were considered in this study. In future research, unmarked reformulation structures should be included to confirm whether reformulation structures (marked and unmarked) are more frequent in source than in target texts.

The second hypothesis, namely the use of signs and gestures of the source signers by the interpreters, is not supported in this dataset. This may be explained by the experimental setting, as there was a camera in front of the interpreters instead of a user. After the recordings, interpreters acknowledged that they tried to control their amount of gesturing by holding their hands together most of the time because they were taught to do so in their training. Still, they produced some pragmatic gestures (Kendon, 2004) and self-adapters—namely touching their face, body, and hands to maintain mental focus and control stress (Ekman & Friesen, 1969). Self-adapters are not included in this paper because they are not used to signal description or depiction in the reformulation structures. The articulators used by interpreters to signal the two modes are the voice, head movements, and facial expressions, showing that interpreters draw on a combination of semiotic resources to construct meaning.

The main channels of expression in both modalities (i.e., the hands in LSFB and the voice in French) are always engaged for description in reformulation structures of the source and target discourses. Description is found in all reformulation structures in both datasets. By contrast, depiction is used in 12 reformulation structures in LSFB, but only in two reformulation structures in French. Interpreters may have changed the mode of signaling unconsciously because of interpreting constraints, or they may have done it deliberately as the opposite strategy to *role shifting* (Heyerick, 2021, p. 125). *Role shifting* is used in spoken-to-signed language interpreting when the interpreter enacts the actions or events in the target text presented from a narrator's point of view in the source text. In the opposite direction, it seems that I002 and I006 use *indirect reporting*, meaning that enactment in LSFB is transformed into indirect speech in French.

Despite the differences in the main channel of expression, the setting (dialogues vs. interpreted renditions), and the type of data (semi-spontaneous vs. interpreted data), the form of the most frequent reformulation structures is <X1> M1 <Y1> in both datasets. Combined or chained reformulation structures, i.e. <X1> M1 <Y1 X2> M2 <Y2>, as well as embedded reformulation structures, i.e., <X1> M1 <Y1 <X2> M2 <Y2> Y1>, are also found in source and target discourses. Although there is a generalized preference for the marker to be placed between the X and the Y segment in the dataset, it is embedded in the Y segment or appears at the end of it in some cases. Furthermore, there is a variety of markers used in LSFB and French expressed through the manual and vocal channels respectively.

The specificity of reformulation structures in target French data is that they can either be interlingual (i.e., generated by the source LSFB signers and reproduced by interpreters in their renditions) or intralingual (i.e., only generated in the interpreted rendition). Hence, these two types of reformulation structures trigger different interpreting strategies. When interlingual reformulation structures are produced, interpreters employ the following strategies (Heyerick, 2021, p. 191):

- Substitution: replacing an item of the source text with something similar but not exactly equivalent in the target text such as a synonym, superordinate, hyponym, or a reference to a locus.
- Omission: eluding information from the source text in the target text.
- Compression: reducing the message of the source text but preserving its meaning.

Since intralingual reformulation structures are not transferred from the source text, they are interpreting strategies *per se* used by interpreters to elaborate their discourse. In other words, intralingual reformulation structures can be seen as the hypernym of the following interpreting strategies (Heyerick, 2021, p. 191):

- Addition: introducing information in the target text which was not present in the source text.
- Paraphrase: using several different signs and/or constructions to present the information from the source text into a longer utterance in the target text.
- Repair: correcting an interpreting mistake, rendering initially omitted information, or improving the initial rendition through an alternative formulation.
- Repetition: giving information, which only appears once in the source text, at least twice in the target text.

Both I002 and I006 produced interlingual and intralingual reformulation structures in their renditions. Sometimes reformulation structures were used for the same chunk of source dialogues, and sometimes not. In other words, interpreters can interpret a marked reformulation of the source dialogue as a marked reformulation in the target discourse (interlingual reformulation), but they can also interpret the structure otherwise or even omit it. Although intralingual reformulations are created on the interpreter's initiative, interpreters can coincide in the chunks where these structures are employed (e.g., to clarify a concept so that the audience may understand it better).

5. Conclusions and future avenues for research

This paper describes reformulation structures in LSFB-to-French interpreting, including their frequency of use, form, and semiotic composition. Reformulations in target renditions can be interlingual or intralingual, depending on whether they appear in the source text or are only created in the target text. Therefore, the phenomenon was analyzed in two datasets of source LSFB and target French data to disentangle the two types of reformulations. The source LSFB data include two dialogues between two signers (totaling 10 minutes) from the LSFB Corpus (Meurant, 2015) and the target French data comprise the renditions of two experienced interpreters (totaling 20 minutes) from the CorMILS Pilot Project (Gabarró-López, 2018). Although reformulation structures are found in both datasets and often exhibit similar forms, they differ in the articulators used to express them.

It was expected that interpreters would rely on signs or gestures articulated by the source signers to produce reformulation structures, as reported in the literature (Janzen et al., 2016). However, this hypothesis was not supported, suggesting new avenues of research. First, the role of self-adapters (Ekman & Friesen, 1969) and embedded gestures, e.g., finger-lift movements while fingers of both hands are in contact (Cienki, 2021, 2023, this volume), should be investigated. These two categories appear several times in the dataset, within reformulation structures and outside them, and may be preferred by interpreters over other categories such as referential gestures, i.e., iconic or deictic gestures that refer to an object, person, location, or event (McNeill, 1992). Second, the renditions of more interpreters should be analyzed. In doing so, the differences in the number and types of gestures could be studied to determine

gestural styles (Zagar Galvão, 2020). Third, nonmanual gestures produced by interpreters when working from the signed-to-spoken direction could also be examined, as nonmanual gestures seem to play a prominent role in conveying meaning that remains unresearched to date.

The main shortcoming of the present research is the small size of the two datasets, which does not allow for broader generalizations. As mentioned earlier, future research should involve more interpreters and more interpreted discourses, ideally renditions that were elicited not only in experimental conditions (i.e., at least the users of the interpreting service should be present). Furthermore, results should be put into perspective with the number of signers producing the source vs. the number of interpreters producing the target. The interpreters' renditions may have been different if each interpreter were interpreting one signer at a time (i.e., the source text was a monologue), or if there were two interpreters in the setting, one for each signer participating in the dialogue.

Despite these shortcomings, this paper provided valuable insights into the use of reformulation structures in LSFB > French interpreting and, more generally, contributed to broadening our knowledge of the signed-to-spoken interpreting direction. This interpreting direction is understudied as compared to the spoken-to-signed interpreting direction (Wang, 2021). Yet, the former may have implications beyond the interpreter's role as a mediator between signers and speakers. In a society where most people lack signing skills and many prejudices surround deaf people, the interpreter's performance (who is voicing the signer's discourse) may influence the judgments of the hearing audience (Feyne, 2015). Therefore, more research on signed-to-spoken language interpreting is needed to provide interpreters with research-based insights. Hopefully, this type of research will see the light of day soon.

6. Acknowledgments

This research was supported by (1) a María Zambrano grant funded by the European Union-NextGenerationEU, Ministry of Universities and Recovery, Transformation and Resilience Plan, through a call from Pompeu Fabra University, and (2) a Postdoctoral Researcher grant funded by the F.R.S. – FNRS (reference 1.B.181.22). I want to thank the participants of the LSFB Corpus and the CorMILS Pilot Project without whom this research would not be possible. I am also grateful to Laurence Meurant for her support with the annotation and her valuable insights, and to the two reviewers whose comments improved the quality of this paper. Any remaining mistakes are my responsibility.

7. References

Blakemore, D. (1993). The relevance of reformulations. *Language and Literature, 2*(2), 101–120. https://doi.org/10.1177/096394709300200202

Bø, V. (In press). Investigating interpreter-mediated interaction through the lens of depictions, descriptions, and indications: An ecological approach. *Translation and Interpreting Studies (TIS)*.

Cienki, A. (2021). From the finger lift to the palm-up open hand when presenting a point: A methodological exploration of forms and functions. *Languages and Modalities, 1*, 17–30. https://doi.org/10.3897/lamo.1.68914

Cienki, A. (2023). Self-adapters: Renewed interest in an often ignored category. *First International Multimodal Communication Symposium (MMSYM 2023)*. 26th – 28th April, Barcelona.

Clark, H. H. (1996). *Using language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Cuenca, M. J., & Bach, C. (2007). Contrasting the form and use of reformulation markers. *Discourse Studies, 9*(2), 149–175. https://doi.org/10.1177/1461445607075347

Cuxac, C. 2007. Une manière de reformuler en langue des signes française. *La linguistique, 43*, 117–128. https://doi.org/10.3917/ling.431.0117

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica, 1*, 49 – 98.

Ferrara, L., & Hodge, G. (2018). Language as description, indication, and depiction. *Frontiers in Psychology, 9*(716), 1–15. https://doi.org/10.3389/fpsyg.2018.00716

Feyne, S. (2015). Typology of interpreter-mediated discourse that affects perceptions of the identity of deaf professionals. In B. C. Nicodemus & K. Cagle (Eds.), *Signed language interpretation and translation research: Selected papers from the First International Symposium* (pp. 49–70). Gallaudet University Press. https://doi.org/10.2307/j.ctv2rh2b69.6

Gabarró-López, S. (2018). *CorMILS: Pilot Multimodal Corpus of French – French Belgian Sign Language (LSFB) interpreters*. Institutionen för lingvistik, Stockholms universitet, Sweden, and LSFB-Lab, Université de Namur, Belgium.

Heyerick, I. (2021). *A descriptive study of linguistic interpreting strategies in Dutch-Flemish Sign Language interpreting. Exploring interpreters' perspectives to understand the what, how and why*. PhD thesis, KU Leuven.

Jakobson, R. (1963). *Essais de linguistique générale*. Éditions de Minuit.

Janzen, T. D., Shaffer, B., & Leeson, L. (2016). Do interpreters draw meaning from a speaker's multimodal text? *7th Conference of the International Society for Gesture Studies (ISGS7)*, Paris. July 18–22.

Johnston, T. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics, 15*(1), 106–131. https://doi.org/10.1075/ijcl.15.1.05joh

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers, 41*(3), 841–849. https://doi.org/10.3758/BRM.41.3.841

McNeill, D. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Meurant, L. (2015). *Corpus LSFB. First digital open access corpus of movies and annotations of French Belgian Sign Language (LSFB)*. Université de Namur, LSFB-Lab. Available at http://www.corpus-lsfb.be

Meurant, L. (2022). Put another way. Reformulation as a window into discourse and interaction in LSFB (French Belgian Sign Language). *Belgian Journal of Linguistics, 36*(1), 145–178. https://doi.org/10.1075/bjl.00074.meu

Meurant, L., Lepeut, A., Tavier, A., Vandenitte, S., Lombart, C., Gabarró-López, S. & Sinte, A. (Ongoing). *The multimodal FRAPé Corpus: Towards building a comparable LSFB and Belgian French corpus*. LSFB-Lab, Université de Namur, Belgium.

Meurant, L., Sinte, A. & Gabarró-López, S. (2022). A multimodal approach to reformulation. Contrastive study of French and French Belgian Sign Language through the productions of speakers, signers and interpreters. *Languages in Contrast, 22*(2), 322–360. https://doi.org/10.1075/lic.00025.meu

Murillo, S. (2016). Sobre la reformulación y sus marcadores. *Cuadernos AISPI: Estudios de lenguas y literaturas hispánicas, 8*, 237–258. https://doi.org/10.14672/8.2016.1335

Peirce, C. S. (1955). *Philosophical writings of Peirce*. Dover Publications.

Pons Bordería, S. (2013). Un solo tipo de reformulación. *Cuadernos AISPI: Estudios de lenguas y literaturas hispánicas, 2*, 151–169. https://doi.org/10.14672/2.2013.1068

Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge University Press.

Wang, J. (2021). *Simultaneous interpreting from a signed language into a spoken language. Quality, cognitive overload, and strategies*. Routledge.

Woroch, J. (2010). *La reformulation comme fondement de l'interprétation de conférence*. PhD thesis, Uniwersytet im. Adama Mickiewicza, Poznań.

Zagar Galvão, E. (2020). Gesture functions and gestural style in simultaneous interpreting. In H.Salaets & G. Brône (Eds.), *Linking up with video: Perspectives on interpreting practice and research* (pp. 151–179). John Benjamins Publishing Company. https://doi.org/10.1075/btl.149.07gal

Sílvia Gabarró-López
Université de Namur
Rue de Bruxelles 61
B-5000 Namur
Belgium

silvia.gabarro@unamur.be

Biography: Sílvia Gabarró-López is a postdoctoral researcher at the University of Namur (Belgium), from which she received her PhD in 2017. Her thesis is the first to study discourse markers from a comparative corpus-based perspective in two sign languages. Before holding her current position, she taught different subjects in the bachelor's and master's degrees in Sign Language Interpreting at the University Saint-Louis-Brussels and UCLouvain (Belgium). She has also worked as a postdoctoral researcher at Stockholm University (Sweden) and Pompeu Fabra University (Spain). She is mainly interested in visual and tactile sign languages, translation & interpreting, contrastive linguistics, discourse analysis, and multimodality.