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.

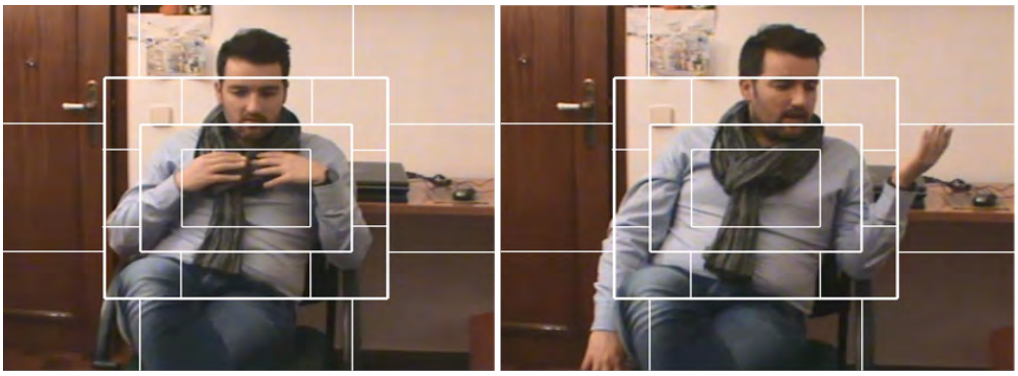


Figure 1. Division of gesture space following McNeill (1992)

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

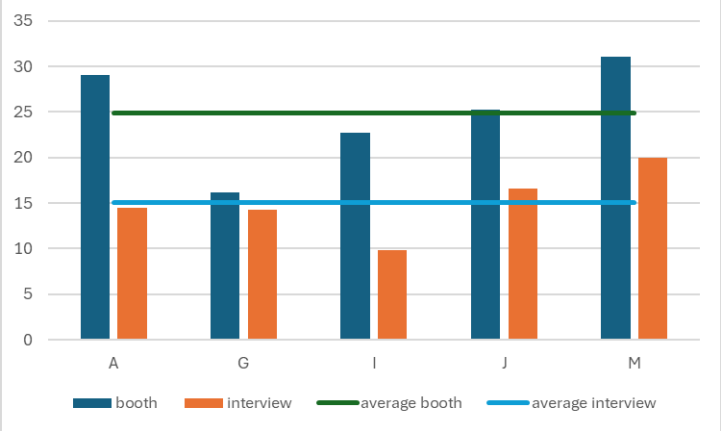


Figure 2. GPM in the booth and in the interview

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

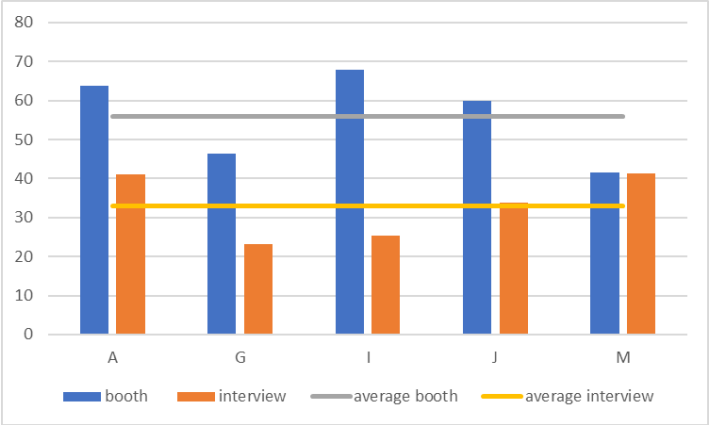


Figure 3. Percentages of small gestures per participant and setting

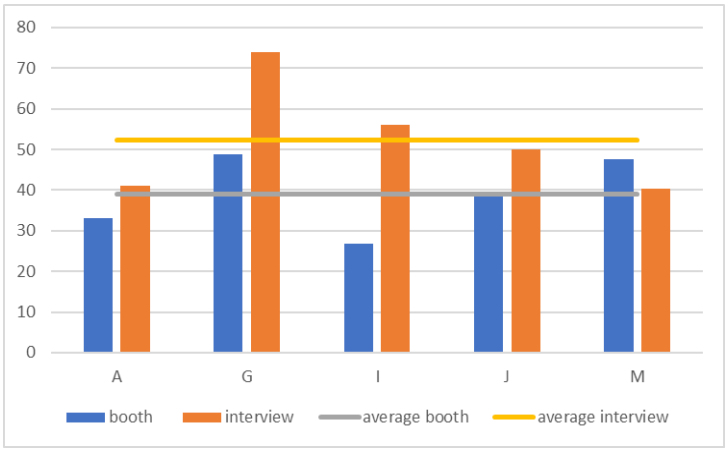


Figure 4. Percentages of medium gestures per participant and setting

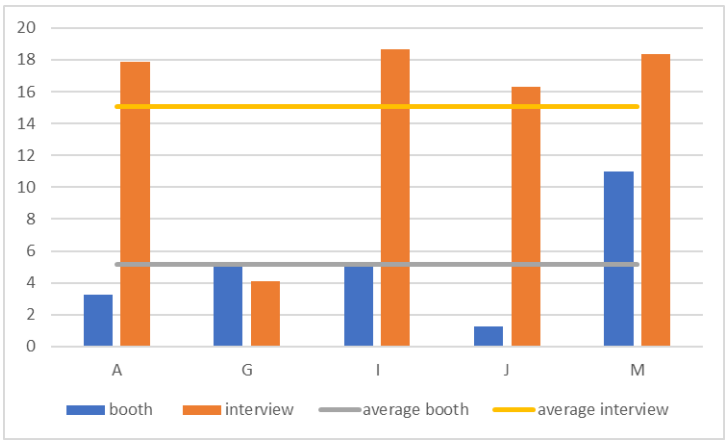


Figure 5. Percentages of large gestures per participant and setting

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

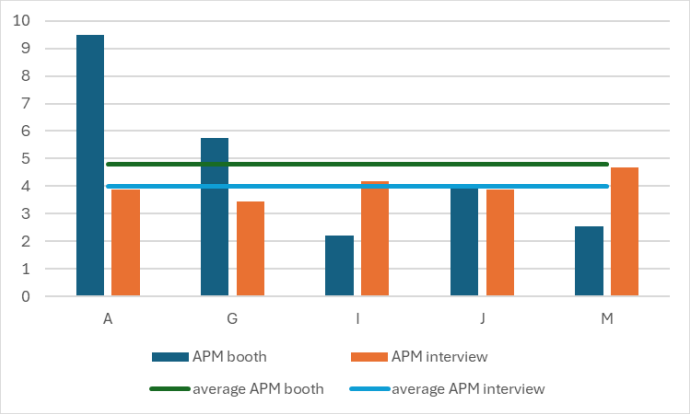


Figure 9. APM in the booth and in the interview

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

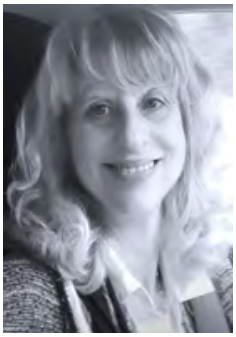
Zagar Galvão, E., & Galhano-Rodrigues, I. (2010) The importance of listening with one's eyes: A case study of multimodality in simultaneous interpreting. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds), *New insights into audiovisual translation and media accessibility* (pp. 239–253). Rodopi. https://doi.org/https://doi.org/10.1163/9789042031814_018

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).