

Article

Gestural Alignment in Spoken Simultaneous Interpreting: A Mixed-Methods Approach

Inés Olza 

Institute for Culture and Society (ICS), University of Navarra, 31009 Pamplona, Spain; iolzamor@unav.es

Abstract: Cognitive and behavioral alignment plays a major role in simultaneous interpreting, the interpreter centrally monitoring and accommodating his/her behavior to that of the speaker-source. In parallel, the place of gesture in the interpreters' practice, as well as its degree of convergence with respect to the gestures of the speaker-source, has been scarcely analyzed until very recently. The multimodal data for this study were collected under (quasi-)experimental conditions in a real court interpreting setting during spoken training exercises performed by two novice interpreters. In this exploratory study, the gestural performance of the interpreters, including their degree of gestural alignment towards the speaker-source, is analyzed and compared using a mixed-methods approach to a randomized sample of the recorded data. The analysis combines a basic descriptive quantification of body movements and a qualitative and comparative analysis of the gesture types performed by the speaker-source and the interpreters. The results show that, in spite of individual differences in interpreting fluency and gestural styles, both interpreters tend to align with the speaker-source's gestural behavior in several ways, and thus a basic taxonomy of gestural convergence between the speaker-source and the interpreters is defined according to several criteria (mainly, gesture presence and gesture type). Our conclusions also allow us to formulate new research questions and hypotheses to be tested in future studies (e.g., types of gestures by the speaker-source that prompt a higher degree of alignment).

Keywords: gesture; alignment; spoken simultaneous interpreting; multimodal data



Citation: Olza, Inés. 2024. Gestural Alignment in Spoken Simultaneous Interpreting: A Mixed-Methods Approach. *Languages* 9: 151. <https://doi.org/10.3390/languages9040151>

Academic Editors: Ana M. Cestero Mancera and Mercedes Díez Prados

Received: 19 February 2024

Revised: 5 April 2024

Accepted: 10 April 2024

Published: 21 April 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Within the realm of non-verbal communication, gesture is nowadays a major focus for interdisciplinary research, with approaches ranging from linguistics to semiotics, cognitive science, psychology, anthropology, or computational science (Müller et al. 2013). All these disciplines agree on the role and potential of gesture both as a 'window' into human cognition (see McNeill 1992, 2005, among others) and a key component of face-to-face communication, the latter understood as one of the most complex forms of interpersonal semiotic behavior (see the seminal works by Poyatos 1994 or Kendon 2004, among many others). Within linguistics, gesture is a multi-faceted object of study. On the one hand, it adds to the fundamental body of research on meaning-making processes (semantics and semiotics of language), which now goes beyond the verbal materiality, paying attention to the multimodal signals and meaning(s) that can be effectively recognized by individuals and/or cultural groups (Calbris 2011; Mittelberg and Hinnell 2022). On the other hand, cognitive linguistics has also turned its attention to gesture as both a reflect and a shaper of human thought, with tight connections with other well-known mechanisms (e.g., metaphor, metonymy) that motivate and pervade language use (McNeill 1992, 2005; Cienki and Müller 2008; Cienki 2022). Both macro-approaches to gesture (the semantic–semiotic one, and the cognitive one) are naturally intertwined. In this paper, however, the second one is more prominent: gesture performance is taken as a material anchor to explore a high-scale cognitive operation—alignment—that unfolds in diverse kinds of human behavior, including language.

The first systematic approach to what was later on called *alignment* was developed by the communication accommodation theory (CAT) (Giles et al. 1991; Giles and Ogay 2007) in the frame of sociolinguistics and social psychology. CAT basically claims that human communication is regulated by constant dynamics of convergence and divergence towards the other(s), these ‘movements’ being materialized in changes and adaptations (accommodation) of the communicative behavior of the speakers at different levels (verbal, paraverbal, non-verbal) (Giles and Ogay 2007, p. 295). In other words, while engaging in linguistic interaction, speakers monitor their own behavior and that of their interlocutors, and consequently—we could even say, ‘strategically’—approach (align) or separate (misalign) their subsequent behavior with respect to that of the other(s).

The notion of ‘accommodation’ in communication was mirrored and expanded into the wider concept of ‘alignment’ by cognitive science and behavioral studies. Thus, (mis)alignment encompasses the convergence–divergence movements between speakers not only in communication (e.g., face-to-face interaction) but also in any other kind of human behavior that involves interaction and cooperation between individuals: e.g., joint actions ranging from physical manipulation of objects (e.g., cooking together) to symbolic tasks (e.g., playing together). Within this cognitive and behavioral frame, and in broad terms, alignment has been defined and analyzed in two distinct manners.

(a) Alignment defined within priming models. Grounded in cognitive psychology (Pickering and Garrod 2004; Menenti et al. 2012), the priming approaches to human behavior—including communication in its diverse forms—stress and demonstrate that participants in communicative interaction 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). Here, alignment is understood as a material manifestation of the wider priming principle that regulates human interaction, that is “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 mental and situational models (Pickering and Garrod 2004, sct. 2.1–2.3). Although priming factors have been widely attested in cases of verbal alignment at different levels (e.g., syntactic alignment, Cleland and Pickering 2003; phonetic alignment, Berry and Ernestus 2018), there is still scarce literature on how the individuals gesturally align in face-to-face communication, nor on gestural alignment in forms of human interaction that do not necessarily involve dialogue—for instance, body postures and movements which are not voluntary and/or are not coordinated with speech (e.g., face rubbing; see the ‘chameleon effect’ described by Chartrand and Bargh 1999). This paper therefore aims to offer exploratory evidence of gestural alignment in a naturally (and highly) ‘chameleonic’ task like simultaneous interpreting.

(b) Alignment defined within grounding models, which have tried to go beyond the priming approaches to human interaction, as they consider the latter to hold a rather mechanist, determinist and/or rigid view of how speakers interact within a stimuli–response dynamics (Krauss and Pardo 2004; Doyle and Franck 2016). While acknowledging imitation as one of the potential ingredients of human interaction, these new models claim that alignment should be rather understood under the scope of interpersonal synergy (Fusaroli and Tylén 2016), where the subjects establish a common ground of understanding about the structure and goals of interaction (Riordan et al. 2014). Here, alignment is understood as a form of synchronized activity which is negotiated in a relational—but not necessarily conscious—way, with wider room for the joint attention and the co-operative action (Eilan et al. 2005; Goodwin 2018) that characterize any form of human conversation. Within this framework, alignment strongly relies on the common goals and communicative dynamics established between the interlocutors in concrete, genre-based, and situated interaction, as is often described by conversational analytic approaches (Stivers 2008; Stivers et al. 2011). The more appropriate understanding of alignment in cooperative—rather than imitative—terms has been also underscored by predictive models drawn from experimental research

(Riordan et al. 2014, pp. 475–77; Fusaroli and Tylén 2016, sct. 6.2 and 6.3; Doyle and Franck 2016, sct. 6).

Bringing the study of alignment into rather uncharted territories, the exploratory study presented in the following sections tackles visibly aligned behavior, as well as the described priming–grounding dialectics (see esp. Section 4.2), during a very peculiar type of communicative ‘interaction’: simultaneous interpreting, where alignment—i.e., behavioral coupling, imitation, and activation of a common ground of understanding with the ‘interlocutor’—is unfolded mainly by the interpreter. Furthermore, as pointed out above, we seek to provide exploratory evidence of gestural alignment, which remains understudied compared to alignment on other levels of linguistic interaction, with most of the existing studies on alignment in gesture focusing on data collected in lab settings (Kimbara 2006; Bergmann and Kopp 2012; Kopp and Bergmann 2013; Rasenberg et al. 2020).

On the other hand, our study aims to contribute to the growing body of research on the presence and functions of gesture in spoken simultaneous interpreting¹, from the perspective of the interpreter him/herself, but also in relation with the original performance of the speaker-source. To our knowledge, empirical research within this subfield has focused thus far on the role of gesture within the overall performance of the interpreters and the cognitive load involved in it (Zagar Galvão 2015, 2020; Stachowiak-Szymczak 2019; Iriskhanova 2020; Martín de León and Fernández Santana 2021; Fernández Santana and Martín de León 2022; Iriskhanova et al. 2023), with very scarce attention to the degree of gestural convergence between the speaker-source and the interpreter, which has been thus far addressed in only a few qualitative micro-analyses (Zagar Galvão 2013). The study presented here tries to overcome this research gap by offering more systematic—and comparative—evidence on the level of gestural alignment exhibited by two different interpreters working at the same time and setting with the same speaker-source.

Against this backdrop, the main objective of this study is to design and apply a mixed-methods analytical model to analyze gestural alignment in spoken interpreting data multimodally recorded in a natural professional setting. The model aims to quantify (in a basic descriptive way), qualify, and compare the degree of gestural alignment towards the same speaker-source exhibited by two distinct interpreters, who were recorded while working at the same time in the mentioned setting. The nature of this study is exploratory, as our analysis seeks to offer a first systematic approach to our data (see Sections 2 and 3) that may allow to formulate and discuss further hypotheses to be statistically tested in future research (see Section 4).

2. Materials and Methods

2.1. The Data

The data were obtained at real training sessions for novice legal interpreters organized by the interpretation directorate of an official international court². A complete ca. 30 min 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.

The speaker-source delivered a speech in Spanish on non-legal issues related to the history of technology³. Two novice interpreters were recorded: Interpreter 1 (female) worked from Spanish into spoken English; and Interpreter 2 (female) worked from Spanish into spoken French. Both interpreters held a similar view trajectory over the speaker, as shown below in Figures 1 and 2. Three cameras were situated next to the main speaker (on his right) and both recorded interpreters (on their right, too). The cameras did not interfere with nor block the activity and view trajectories of the participants.



Figure 1. General overview of the recording setting (real courtroom).

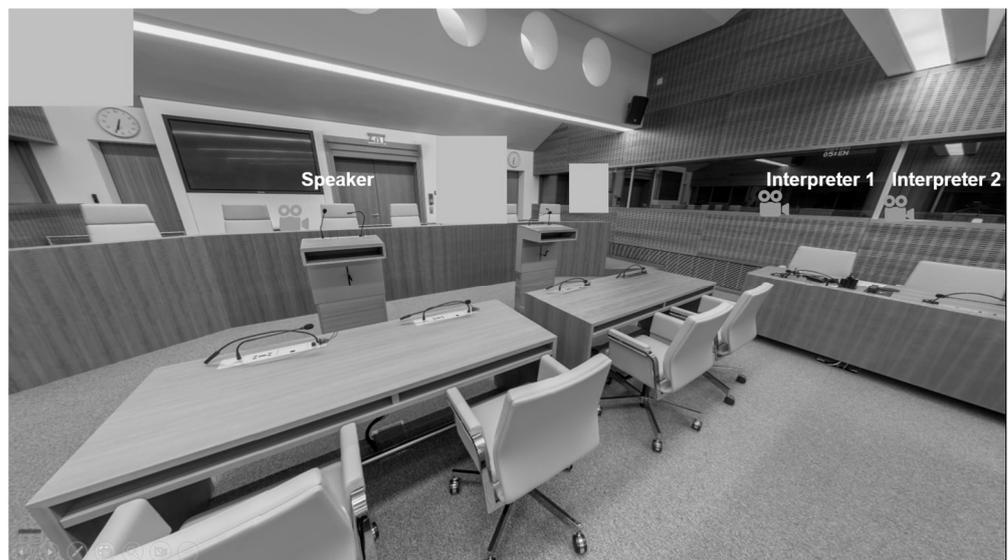


Figure 2. Spatial distribution: speaker-source, recorded interpreters, and video-cameras.

2.2. Research Questions, Exploratory Hypotheses, and Mixed-Methods Analysis

As explained in Section 1, several strands of empirical research have shown that participants engaging in communicative interaction tend to align—behaviorally converge—at several linguistic levels, gesture being the one which has thus far received less attention (Rasenberg et al. 2020). At the same time, gesture research in simultaneous interpreting has decidedly increased in recent years, but the connections between the interpreter and his/her speaker-source at the gestural level are still to be systematically explored (Zagar Galvão 2013; Chwalczuk 2021).

In light of all this, the mixed-methods approach used to analyze the data in this exploratory study was informed by the following research questions and their corresponding exploratory hypotheses, which jointly aim at bridging the described gaps in the study of gestural alignment in simultaneous interpreting.

- Research Question 1. Do the simultaneous interpreters gesturally align with the speaker-source, that is, are there clear examples of gestural alignment by the two interpreters?

- Exploratory Hypothesis 1. Cases of gestural alignment by both interpreters are found in the data.
- Research Question 2. Are there relevant differences in the degree of gestural alignment exhibited by the two interpreters?
 - Exploratory Hypothesis 2. Although some differences may arise between them (due to individual styles and/or fluency), both interpreters show a similar—or reasonably similar—degree of gestural alignment towards the speaker-source, alignment being a general high-scale cognitive principle operating in any simultaneous interpreter’s performance.
- Research Question 3. Do certain gesture types prompt a higher degree of alignment?
 - Exploratory Hypothesis 3. Because of their semantically oriented content (representation of objects, subjects, facts, etc.), representational gestures by the speaker-source (e.g., iconic or metaphoric gestures) are more likely to be replicated by the interpreters.

The three multimodal recording files (speaker-source, Interpreter 1, Interpreter 2) were analyzed and tagged separately using the annotation tool ELAN-6.5 (<https://archive.mpi.nl/tla/elan>; accessed on 15 August 2023). The analysis was run by a single coder (the author of this paper) according to the following steps.

Step 1. Analysis of four 1 min excerpts of the speaker-source’s behavior, which were later on used as a baseline for the comparative analysis of the performance of Interpreters 1 and 2 (Step 2). The excerpts were chosen randomly using an open-source aleatory choice generator (<https://randomchoicegenerator.com/>; accessed on 23 March 2023), resulting in minutes 2:00–3:00; 10:00–11:00, 20:00–21:00, and 27:00–28:00. The selected excerpts were qualitatively analyzed in ELAN. First, the presence of any gesture relevant to the speech content was annotated. At this point, beats⁴ and self-adaptors⁵ were excluded for two reasons: they tend to rely more strongly on the style of each speaker (they may be thus subject to a more dramatic individual variation) and they do not relate with what is signaled or represented by speech. 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], as reflected in Figure 3. The following gesture types were distinguished and coded (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 (e.g., enumerations); modality and stance—gestures for intensification or mitigation of the expressed content; and gestures for negation. Following Kendon’s (2004) broader distinction, iconic and metaphoric gestures are representational gestures, insofar as they refer to the utterance content; and discourse and information gestures and modality and stance gestures, as well as gestures for negation, are pragmatic gestures, as they refer to the utterance itself, that its, to its enunciative and structural properties.

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 1.

Table 1. Speaker-source: total number of gestures and gesture types.

Speaker-Source		
	Hits	Rate
Discourse and information structure	29	42.6%
Iconic	18	26.5%
Metaphoric	8	11.8%
Modality and stance	7	10.3%
Negation	6	8.8%
Total	68	100%

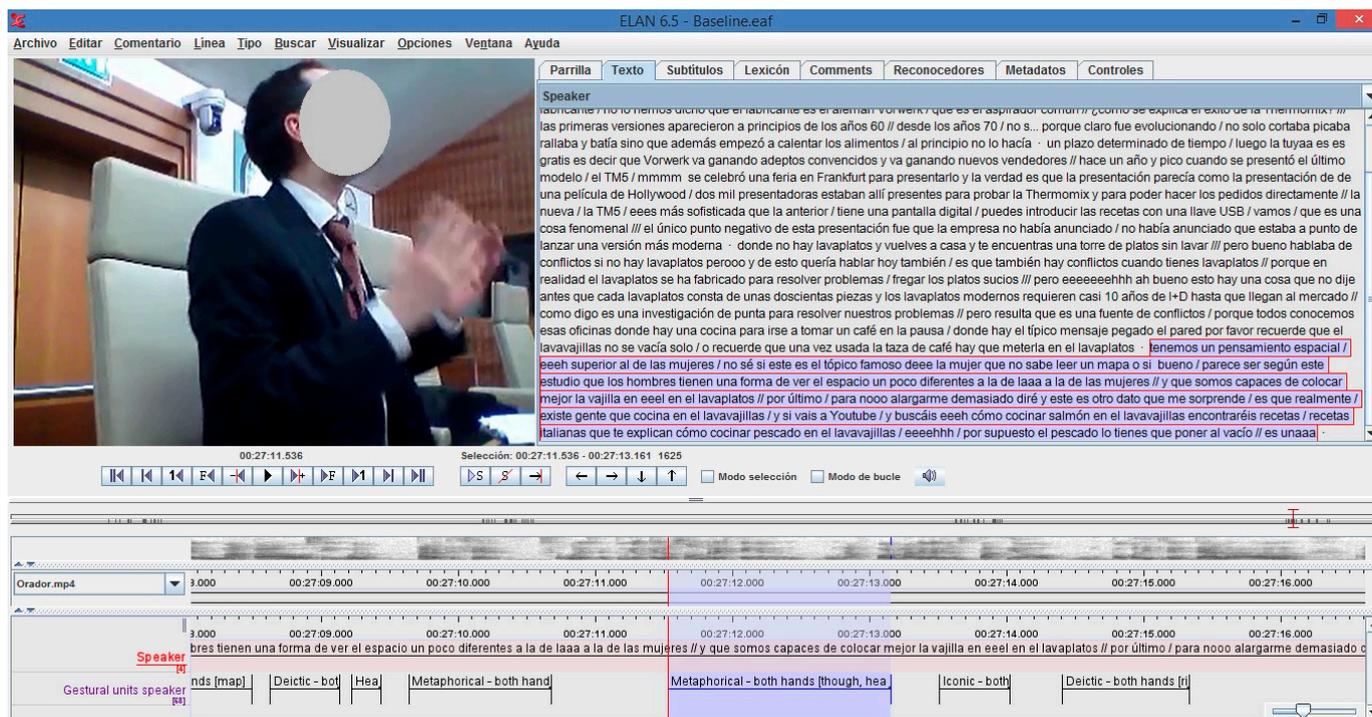


Figure 3. Annotation of the speaker-source's gestures in ELAN.

Step 2. Analysis of the interpreters' performance in the corresponding excerpts of their videos—excerpts where the professionals interpret the speech uttered by the speaker-source in the minutes analyzed in Step 1. For Interpreter 1, minutes 6:07–7:20, 14:20–15:20, 24:10–25:10, and 31:00–32:10 were analyzed; for Interpreter 2, minutes 4:05–5:20, 12:10–13:10, 22:00–23:00, and 29:00–30:00 were examined. As we said before, the speaker-source's behavior was taken as a baseline to analyze the interpreters' performance. Therefore, the sequences where the speaker-source gestured—verbal cue + relevant gesture going along with it—were contrasted with the corresponding interpretation by both professionals. Their behavior in the corresponding sequences was qualitatively analyzed and annotated using the following tags.

- 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; consequently, there is no chance for the interpreter to (mis)align with the gestures 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 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—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.

Again, a basic descriptive quantification of the interpreters' performance and a rough quantitative comparison of their behavior and that of the speaker are offered in Section 3 (see Tables 2–4).

Table 2. Interpreters’ overall gesture performance.

	Interpreter 1		Interpreter 2	
	Hits	Rate	Hits	Rate
Speech interpreted—same type of gesture	24	35.3%	15	22.05%
Speech interpreted—different type of gesture	21	30.9%	21	30.9%
Speech interpreted—no gesture	5	7.3%	17	25%
Speech not interpreted	18	26.5%	15	22.05%
Speaker-source baseline	68	100%	68	100%

Table 3. Speech + gesture hits by the speaker-source that are interpreted along with a gesture.

	Baseline: Gestures by Speaker-Source	Interpreter 1: Nr of Hits Where Speech Is Interpreted Along with a Gesture		Interpreter 2: Nr of Hits Where Speech Is Interpreted Along with a Gesture	
	Hits	Hits	Rate *	Hits	Rate *
Discourse and inf structure	29	18	62%	14	48.2%
Iconic	18	14	77.9%	11	61.1%
Metaphoric	8	5	62.5%	1	12.5%
Modality and stance	7	7	100%	7	100%
Negation	6	1	16.6%	3	50%
	68	45		36	

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

Table 4. Speech + gesture hits by the speaker-source that are interpreted along with a gesture of the same kind.

	Baseline: Gestures by Speaker-Source	Same Gesture Type by Interpreter 1		Same Gesture Type by Interpreter 2	
	Hits	Hits	Rate *	Hits	Rate *
Discourse and inf structure	29	10	34.5%	5	17.2%
Iconic	18	10	55.5%	7	38.9%
Metaphoric	8	0	0%	0	0%
Modality and stance	7	4	57.1%	3	42.8%
Negation	6	0	0%	0	0%
	68	24		15	

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

3. Results

3.1. Gestural Behavior of the Speaker-Source in the Selected Excerpts

Within the aleatory excerpts selected as a baseline for the speaker-source’s performance, the quantitative outbreak per gesture type (see Table 1) reveals a major proportion of gestures carrying out discourse structuring functions (e.g., bringing up new topics or pointing to already mentioned facts or entities). Also, as expected (the speech dealt with the history of technology), representative gestures—either iconic or metaphoric—are a second major type in the speaker-source’s performance.

3.2. Gestural Behavior of the Interpreters with Respect to That of the Speaker-Source

The research questions presented in Section 2.2 were addressed when quantifying and relating the interpreters’ behavior to that of the speaker-source.

Research Questions 1 and 2 were as follows: Do the simultaneous interpreters gesturally align with the speaker-source, that is, are there clear examples of gestural alignment

by the two interpreters? Are there relevant differences in the degree of gestural alignment exhibited by the two interpreters? To jointly answer these questions, the overall interpreters' performance with regard to the speaker-source baseline was quantified, which also allowed for a general comparison of the behavior of Interpreter 1 and Interpreter 2.

A basic quantification of the categories explained in Section 2.2 is included in Table 2. As shown there, the interpreters align with the gestural behavior of the speaker-source in a good number of cases, where they carry out the same type of gesture observed in the speaker-source (35.3% of cases for Interpreter 1; 22.05% for Interpreter 2).

Furthermore, there are quite a few cases where the original speech by the speaker-source is not interpreted—the novice interpreters skip that sequence due to disfluencies or gaps in the overall interpreting process. Although their level of inaccuracy in this respect is quite similar (they are trainees at the same stage of their learning program), they differ in their overall degree of gestural alignment: while Interpreter 1 (from Spanish into English) uses the same type of gesture as the speaker-source in 35.3% of the hits (speech interpreted—same type), Interpreter 2 (from Spanish into French) does not even gesture in 25% of the cases where she interprets the speaker-source's speech (speech interpreted—no gesture).

As for Research Question 3 (Do certain gesture types prompt a higher degree of alignment?), first Table 3 reflects the number of speech + gesture hits by the speaker which are effectively interpreted by each interpreter along with a relevant gesture, regardless the specific type of gesture used by the interpreter—e.g., the first row indicates that the 29 hits where the speaker-source used a discourse and information structure gesture prompted an speech + gesture interpretation by the interpreters, their gestures being of any kind (discourse structuring *or* iconic *or* metaphoric and so forth). In broad terms, the performance by Interpreter 1 seems to be more gesturally aligned to the speaker-source than that of Interpreter 2. Modality and stance gestures by the speaker-source were the only type that always prompted a gestural response by both interpreters (7 hits; 100% of the hits by the speaker-source). Iconic gestures also prompted a high degree of gestural response by both interpreters (Interpreter 1, 77.9%; Interpreter 2, 61.1%), followed by discourse structuring gestures (Interpreter 1, 62%; Interpreter 2, 48.1%). The remaining gesture types (metaphoric gestures and gestures for negation) prompted a lower and disparate response between both interpreters.

Table 4 shows the results obtained when narrowing down the scope of results by asking: Which kinds of gestures by the speaker-source more often prompt the *same* type of gestures by the interpreters? As pointed out above, the overall performance of Interpreter 1 aligns more with the speaker-source than that of Interpreter 2. And again, the same types of gesture (modality and stance, and iconic gestures, first; then, discourse structuring gestures) prompt a higher degree of gestural convergence by both interpreters. Metaphoric and negation gestures always prompt gestures of a different kind.

4. Discussion

4.1. Tracking the 'Chameleon Effect': Degrees of Gestural Alignment

The results presented in Section 3 also allow to address and discuss the exploratory hypotheses stated back in Section 2.2.

- Research Question 1. Do the simultaneous interpreters gesturally align with the speaker-source?
 - Exploratory Hypothesis 1. Cases of gestural alignment by both interpreters are found in the data.

The hypothesis is confirmed by the results included above in Tables 3 and 4: not only do the interpreters gesture when working with a multimodal prompt (speech + gesture) by the speaker source, but they sometimes use the same kind of gesture observed in the source. It is therefore possible to draw a progression in the level of behavioral and gestural alignment towards the speaker-source exhibited by the interpreters in our study.

As shown in Figures 4–7, when the interpreters actually tackle and interpret a multi-modal sequence (speech + gesture) produced by the speaker-source, they may:

- Not gesture (gestural misalignment): see Example 1 (Figure 5), where the speaker-source performs an iconic gesture and Interpreter 2 does not gesture;
- Perform a gesture of a different kind compared to that of the speaker-source (intermediate degree of gestural alignment): see Example 2 (Figure 6), where the speaker-source performs a metaphoric gesture that depicts the sellers gained by an enterprise as objects that are caught and brought into a closer space. When working with this sequence, Interpreter 1 also moves her hands, but in a non-representational way—she uses beat gestures to accompany the rhythm of her own speech;
- Perform the same type of gesture as the speaker-source (highest degree of gestural alignment): see Example 3, where Interpreter 1 somehow mimics—in a ‘chameleonic’ way—the iconic wrapping gesture carried out by the speaker-source.

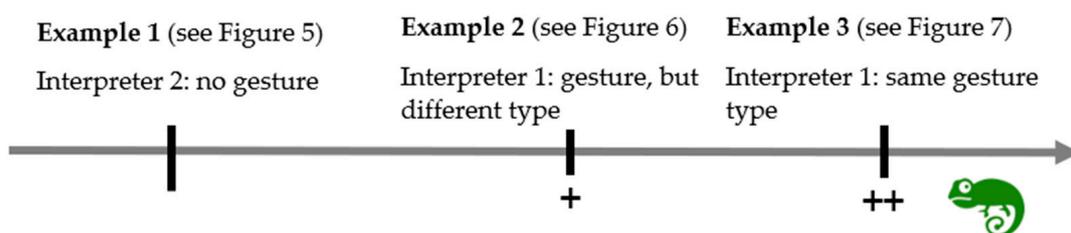


Figure 4. The ‘gestural alignment continuum’.

<p>Speaker-source Iconic gesture, both hands: bringing something to his personal space</p>	
 <p>(a)</p>	<p>luego la tuyaaa es es gratis (a) then yooours is is free</p>
<p>Interpreter 2 No gesture: hands resting at table</p>	
 <p>(b)</p>	<p>le votre sera gratuit (b) then yours will be free</p>

Figure 5. Example 1: no gestural alignment by Interpreter 2.

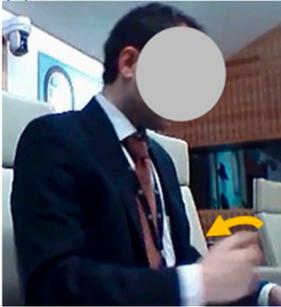
<p>Speaker-source Metaphoric gesture, right hand: bringing something to his personal space</p>	
<p>(a) </p> <p>(b) </p>	<p>va ganando nuevos vendedores (a) (b) she gains new sellers</p>
<p>Interpreter 1 Beat gestures</p>	
<p>(c) first beat </p> <p>(d) second beat </p>	<p>you can recruit new people (c) (d)</p>

Figure 6. Example 2: Interpreter 1 performs a different gesture type.

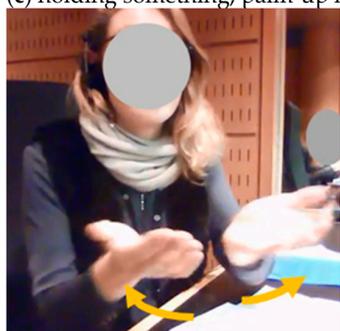
<p>Speaker-source Iconic gesture, both hands: round shape, as if wrapping something</p>	
<p>(a) </p> <p>(b) </p>	<p>el pescado lo tienes (a) (b) que poner al vacío you have to put the fish in vacuum</p>
<p>Interpreter 1 Iconic gesture, both hands: round shape, as if wrapping something</p>	
<p>(c) holding something, palm-up hands </p> <p>(d) holding something from the sides </p>	<p>put the salmon uh (c) to wrap the salmon (d)</p>

Figure 7. Example 3: Interpreter 1 performs the same gesture type as the speaker-source.

This general graduation in gestural alignment can be also applied to compare the performance of both interpreters, thus addressing the second research question of this exploratory study.

- Research Question 2. Are there relevant differences in the degree of gestural alignment exhibited by the two interpreters?
 - Exploratory Hypothesis 2. Although some differences may arise between them (due to individual styles and/or fluency), both interpreters show a similar—or reasonably similar—degree of gestural alignment towards the speaker-source, alignment being a general high-scale cognitive principle operating any simultaneous interpreter’s performance.

The hypothesis was not confirmed. Although the data analyzed in the present study—small and qualitatively addressed data—do not allow the claim for statistically significant differences between both interpreters, bigger divergences than initially expected were still found between them. In particular, the results in the overall performance of both interpreters (see above Table 2) reveal a substantially higher level of gestural alignment for Interpreter 1, who (a) performed the same type of gesture as the speaker-source in a bigger number of cases (35.5% of the interpreted hits versus 22.05% by Interpreter 2); and (b) presented a much lower degree of non-gestural hits, i.e., cases where the interpreted sequence was not accompanied by a gesture (7.3% for Interpreter 1; 25% for Interpreter 2). Given that the overall fluency of their interpretation seemed to be rather similar (as shown by the close percentages of non-interpreted sequences: 26.5%, for Interpreter 1; 22.05% for Interpreter 2), individual styles—e.g., Interpreter 1 perhaps gesturing more often in her general communicative style?—may be responsible for the detected differences, the overarching cognitive and behavioral principle of gestural alignment therefore being more dependent than expected on individual variables. These alternative hypotheses shall be tested in future studies.

4.2. Gesture Types

- Research Question 3. Do certain gesture types prompt a higher degree of alignment?
 - Exploratory Hypothesis 3. Because of their semantically oriented content (representation of objects, subjects, facts, etc.), representational gestures by the speaker-source (e.g., iconic or metaphoric gestures) are more likely to be replicated by the interpreters.

The results obtained in this study can only partially confirm the third exploratory hypothesis. As shown above (see Tables 3 and 4), iconic (representational) and modality-stance (non-representational or pragmatic) gestures prompted gestural alignment more consistently in our data, resulting not only in prompting body movements by the interpreters, but also the performance of same-kind gestures in a larger proportion of hits. In contrast, metaphoric hits (representational) prompted a lower degree of gestural behavior, with no replication of figurative gestures by the interpreters. These results may re-direct our hypothesis towards a different distinction, namely the *priming* versus *grounding* approaches to alignment explained in Section 1. Accordingly, we could hypothesize that gestures prompting a higher gestural alignment may be more ‘easily’—automatically—replicated, the remaining gestural categories reflecting a more complex set of thought subject to a more ‘deliberative’ or ‘strategic’ (non-automatic) response by the interpreters⁶. In other words, we could think that iconic gestures and gestures to express modality and stance (attitudes and emotions of the speaker) are better primers than—for instance—metaphoric gestures, a possible reason for this being a higher degree of conceptual complexity and/or elaboration motivating the figurative mappings of metaphoric gestures. This hypothesis should indeed be further tested in new analyses that put additional analytical categories into play in more systematic ways: namely, conceptual complexity, schematicity in meaning and gesture performance, or concreteness and abstraction in meaning and language.

4.3. Limitations of This Study and New Research Questions for Future Studies

This study is a first exploratory step to understand how gestural alignment works in spoken simultaneous interpreting. However, it has a number of limitations that make it necessary to collect and analyze data in a more thorough manner so as to ensure that an experimentally and statistically valid study can be developed.

A first limitation has to do with the ecological validity of the recorded data as representers of non-biased, authentic interpretation tasks. Although being part of real and situated exercises run in an authentic court setting, the recorded sessions were training exercises for novice interpreters, which means several things:

- (a) The performance of the subjects analyzed may have been compromised by their still incipient expertise;
- (b) The behavior of the speaker-source was documented to be especially cooperative with the interpreters, with a moderate speech speed and use of more frequent and wider/more visible co-speech gestures;
- (c) The physical presence of the trainers and the researcher in the physical setting may have affected the behavior of the novice interpreters.

A second limitation of the study relates to the peculiar interactive nature of any interpreting activity. In this sense, it remains clear that interpreters do *not* strictly *interact* with the speaker-source, as a speaker who engages in face-to-face dialogue with an interlocutor actually does. As most of the evidence on gestural alignment in interaction comes from prototypical forms of dialogue/conversation, future studies will need to address and define the specificities of gestural alignment in a non-dialogical activity like simultaneous interpreting, where the 'interaction' between the subjects is shaped differently—basically, the interpreter is called to reproduce and mediate the speaker-source's behavior.

Finally, several methodological limitations should be pointed out.

- (a) The qualitative analysis was run at all stages by a single coder, which means that at least a second coder should be involved in future studies to grant objectivity and a solid inter-rater reliability. Also, the coder at this exploratory study was not blind to the speaker-source's behavior when analyzing the interpreters' performance. In further analyses, more coders will be involved to ensure that recorded subjects are examined by separate analyzers;
- (b) The present study was carried out through a mixed-methods approach, combining qualitative analysis and basic descriptive quantifications of the observed results. It did not allow for proper statistical analysis of the level of gestural alignment by the interpreters towards the speaker-source, nor of the differences between the interpreters found in the data. To achieve this goal: (b.1) wider samples of the recorded data will be selected and analyzed in forthcoming studies; (b.2) future analyses will comprise all kinds of body movements executed by the speaker-source and the interpreters—i.e., it shall also include beat gestures and self-adaptors, thus avoiding a possible bias introduced by interpretatively informed gestural categories;
- (c) The speaker-source's speech-gesture units were transcribed and/or coded, but the interpreters' speech was not attended in this study. Further studies will necessarily look at the relationship between the interpreters' gestures and their own speech, to understand the concurring reasons why they gesturally align or not with the speaker-source. In the same vein, gesture synchronization with speech will be also considered when addressing the interpreters' performance, to determine whether their gestures synchronize more with the speaker-source's observed behavior or, conversely, with their own speech.

5. Conclusions

The main findings of this exploratory study can be summarized as follows.

- (a) The data collected in a real setting under (quasi-)experimental conditions show that simultaneous interpreters often align with the gestural behavior of the speaker-source,

- with a good number of cases where they replicate the same gesture type observed in the speaker;
- (b) Although showing a similar level of expertise and interpreting accuracy, notable differences were detected in the individual style of the two interpreters and—perhaps consequently—in their degree of gestural convergence towards the speaker-source;
 - (c) In spite of their individual differences, both interpreters aligned more with iconic and modality-stance gestures, and less with discourse structuring gestures, gestures for negation, and metaphoric gestures.

These results should be further explored using wider data samples that are more thoroughly quantified.

Funding: This research was funded by the Spanish Ministry of Science and Innovation, 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).

Institutional Review Board Statement: The study was approved by the Ethics Committee of the University of Navarra (protocol code 2017.021; date of approval: 16 March 2017).

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study. Written informed consent was also obtained from subjects to publish this paper.

Data Availability Statement: Due to ethical restrictions and EU regulation (General Data Protection Regulation, Regulation (EU) 2016/679), the multimodal data analyzed in this study (recordings of interpreting training sessions at an official court) cannot be made publicly available. For controlled access to anonymized data that can be used for replication purposes, please contact the author.

Acknowledgments: Thanks are due to the interpreters who generously agreed to be recorded for this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Notes

- ¹ Major milestones in the field include thus far the defense and/or publication of a (small) number of PhD dissertations in different countries (Portugal: [Zagar Galvão 2015](#); Poland: [Stachowiak-Szymczak 2019](#); France: [Chwalczuk 2021](#)); the publication of the collective volume *Linking up with Video* ([Salaets and Brône 2020](#)), where a ‘multimodal turn’ in interpreting research and training is proposed; and the panel ‘Gesture in spoken and signed-to-spoken language interpreting’, convened by [Silvia 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 pragmatic, cognitive and applied linguistics discussed the latest advancements in gesture analysis in consecutive and simultaneous interpreting of spoken and signed-to-spoken language. An earlier version of this article was presented at this panel—I would like to thank the convenors and participants for their insightful comments and suggestions.
- ² The name and coordinates of the institution are not facilitated due to EU regulation on confidentiality and data protection.
- ³ 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.
- ⁴ Beats are body movements—most usually, hand movements—that go along with the rhythmical pulsation of speech—i.e., prosody, and whose form remains the same regardless of the content expressed in speech ([McNeill 1992](#), p. 15).
- ⁵ Self-adaptors are non-signaling gestures—i.e., they are not intended to convey a particular meaning—where one part of the body is applied to another part of the body, such as scratching one’s head and face, or where one body part moves involuntarily—e.g., tapping the foot ([Ekman and Friesen 1972](#), pp. 362–63; [Koda and Mori 2014](#)).
- ⁶ For the automatic versus strategic functioning of gestural alignment, see also [Kopp and Bergmann \(2013\)](#).

References

- Bergmann, Kirsten, and Stefan Kopp. 2012. Gestural alignment in natural dialogue. *Proceedings of the 34th Annual Conference of the Cognitive Science Society* 34: 1326–31. Available online: <https://escholarship.org/uc/item/73z0q063> (accessed on 10 January 2023).
- Berry, Grant M., and Mirjam Ernestus. 2018. Phonetic alignment in English as a lingua franca: Coming together while splitting apart. *Second Language Research* 34: 343–70. [CrossRef]
- Calbris, Geneviève. 2011. *Elements of Meaning in Gesture*. Amsterdam and Philadelphia: John Benjamins.
- Chartrand, Tanya L., and John A. Bargh. 1999. The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology* 76: 893–910. [CrossRef] [PubMed]
- Chwalczuk, Monika. 2021. La gestualité co-verbale en interprétation dans les services publics: Analyse contextualisée d'un corpus multimodal. Doctoral dissertation, Université Paris Cité, Paris, France. Available online: <https://theses.hal.science/tel-03509754/document> (accessed on 18 September 2023).
- Cienki, Alan J. 2022. The study of gesture in cognitive linguistics: How it could inform and inspire other research in cognitive science. *WIREs Cognitive Science* 13: e1623. [CrossRef] [PubMed]
- Cienki, Alan J., and Cornelia Müller, eds. 2008. *Metaphor and Gesture*. Amsterdam and Philadelphia: John Benjamins.
- Cleland, Alexandra A., and Martin J. Pickering. 2003. The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language* 49: 214–30. [CrossRef]
- Doyle, Gabriel, and Michael C. Franck. 2016. Investigating the sources of linguistic alignment in conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*. Edited by Katrin Erk and Noah A. Smith. Berlin: Association for Computational Linguistics, pp. 526–36. [CrossRef]
- Eilan, Naomi, Christoph Hoerl, Teresa McCormack, and Johannes Roessler, eds. 2005. *Joint Attention: Communication and Other Minds*. Oxford: Oxford University Press.
- Ekman, Paul, and Wallace V. Friesen. 1972. Hand movements. *The Journal of Communication* 22: 353–74. [CrossRef]
- Fernández Santana, Alba, and Celia Martín de León. 2022. Con la voz y las manos: Gestos icónicos en interpretación simultánea. *Hermēneus* 23: 225–71. [CrossRef]
- Fusaroli, Riccardo, and Kristian Tylén. 2016. Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science* 40: 145–71. [CrossRef] [PubMed]
- Giles, Howard, and Tania Ogay. 2007. Communication Accommodation Theory. In *Explaining Communication: Contemporary Theories and Exemplars*. Edited by Bryan B. Whaley and Wendy Samter. London: Routledge, pp. 293–310.
- Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. In *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Edited by Howard Giles, Justine Coupland and Nikolas Coupland. Cambridge: Cambridge University Press, pp. 1–68.
- Goodwin, Charles. 2018. *Co-Operative Action*. Cambridge and New York: Cambridge University Press.
- Healey, Patrick G. T. 2004. Dialogue in the degenerate case? *Behavioral and Brain Sciences* 27: 201. [CrossRef] [PubMed]
- Irishanova, Olga. 2020. *Co-Speech Gestures in Professional Contexts: A Multimodal Analysis of Simultaneous Interpretation*. IMCC Seminar Talk. Oxford: University of Oxford, November. Available online: <https://www.youtube.com/watch?v=87XxMm4e1GU> (accessed on 3 June 2023).
- Irishanova, Olga, Alan Cienki, Maria V. Tomskaya, and Alina I. Nikolayeva. 2023. Silent, but salient: Gestures in simultaneous interpreting. *Research Result. Theoretical and Applied Linguistics* 9: 99–114. [CrossRef]
- Kendon, Adam. 2004. *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kimbara, Irene. 2006. On gestural mimicry. *Gesture* 6: 36–61. [CrossRef]
- Koda, Tomoko, and Yuko Mori. 2014. Effects of an agent's displaying self-adaptors during a serious conversation. In *Intelligent Virtual Agents. IVA 2014*. Edited by Timothy Bickmore, Stacy Marsella and Candace Sidner. Cham: Springer, vol. 8637. [CrossRef]
- Kopp, Stefan, and Kirsten Bergmann. 2013. Automatic and strategic alignment of co-verbal gestures in dialogue. In *Alignment in Communication: Towards a New Theory of Communication*. Edited by Ipke Wachsmuth, Jan de Ruiter, Petra Jaacks and Stefan Kopp. Amsterdam and Philadelphia: John Benjamins, pp. 87–108.
- Krauss, Robert M., and Jennifer S. Pardo. 2004. Is alignment always the result of automatic priming? *Behavioral and Brain Sciences* 27: 203–4. [CrossRef] [PubMed]
- Martín de León, Celia, and Alba Fernández Santana. 2021. Embodied cognition in the booth: Referential and pragmatic gestures in simultaneous interpreting. *Cognitive Linguistic Studies* 8: 277–306. [CrossRef]
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: The University of Chicago Press.
- McNeill, David. 2005. *Gesture and Thought*. Chicago: The University of Chicago Press.
- Menenti, Laura, Martin J. Pickering, and Simon Garrod. 2012. Toward a neural basis of interactive alignment in conversation. *Frontiers in Human Neuroscience* 6: 185. [CrossRef] [PubMed]
- Mittelberg, Irene, and Jennifer Hinnell. 2022. Gesture studies and semiotics. In *Semiotic Movements*. Edited by Jamin Pelkey and Paul Copley. London: Bloomsbury Academic, pp. 183–214.
- Müller, Cornelia, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill, and Sedinha Tessorodorf, eds. 2013. *Body–Language–Communication. An International Handbook on Multimodality in Human Interaction*. Berlin: Mouton de Gruyter, vol. 2.
- Pickering, Martin J., and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27: 169–90. [CrossRef] [PubMed]

- Poyatos, Fernando. 1994. *La Comunicación no Verbal*. Madrid: Istmo, vol. 2.
- Rasenberg, Marlou, Asli Özyürek, and Mark Dingemans. 2020. Alignment in Multimodal Interaction: An Integrative Framework. *Cognitive Science* 44: e12911. [[CrossRef](#)] [[PubMed](#)]
- Riordan, Monica A., Roger J. Kreuz, and Andrew M. Olney. 2014. Alignment is a function of conversational dynamics. *Journal of Language and Social Psychology* 33: 465–81. [[CrossRef](#)]
- Salaets, Heidi, and Geert Brône, eds. 2020. *Linking Up with Video. Perspectives on Interpreting Practice and Research*. Amsterdam and Philadelphia: John Benjamins.
- Stachowiak-Szymczak, Katarzyna. 2019. *Eye Movements and Gestures in Simultaneous and Consecutive Interpreting*. Cham: Springer.
- Stivers, Tanya. 2008. Stance, alignment and affiliation during storytelling: When nodding is a token of preliminary affiliation. *Research on Language in Social Interaction* 41: 29–55. [[CrossRef](#)]
- Stivers, Tanya, Lorenza Mondada, and Jakob Steensig. 2011. Knowledge, morality and affiliation in social interaction. In *The Morality of Knowledge in Conversation*. Edited by Tanya Stivers, Lorenza Mondada and Jakob Steensig. Cambridge: Cambridge University Press, pp. 3–24.
- Zagar Galvão, Elena. 2013. Hand gestures and speech production in the booth: Do simultaneous interpreters imitate the speaker? In *Estudos de Lingüística*. Edited by Conceição Carapinha and Isabel Santos. Coimbra: Coimbra University Press, pp. 115–29.
- Zagar Galvão, Elena. 2015. *Gesture in Simultaneous Interpreting from English into European Portuguese: An Exploratory Study*. Unpublished Doctoral dissertation, University of Porto, Porto, Portugal.
- Zagar Galvão, Elena. 2020. Gesture functions and gestural style in simultaneous interpreting. In *Linking Up with Video. Perspectives on Interpreting Practice and Research*. Edited by Heidi Salaets and Geert Brône. Amsterdam and Philadelphia: John Benjamins, pp. 151–80.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.