The at-issue status of viewpoint gestures: Evidence for gradient at-issueness

Sebastian WALTER — *University of Wuppertal, University of Frankfurt*

Abstract. Recent semantic research on the meaning contribution of speech-accompanying gestures has focused on their at-issue status. Empirical evidence suggests they contribute not-at-issue meaning by default (e.g., Ebert et al., 2020). However, there is growing evidence that the notion of at-issueness is better captured as a gradient one instead of a binary one (Barnes et al., 2022; Tonhauser et al., 2018; see Barnes and Ebert, 2023 for a formal implementation). Research investigating differences in at-issueness between different types of co-speech gestures is missing entirely. This paper presents findings from two experimental rating studies investigating differences in the at-issue status and salience of *character viewpoint gestures* (CVGs) and *observer viewpoint gestures* (OVGs). The results of the first experiment suggest that while both viewpoint gesture types contribute not-at-issue meaning by default, CVGs are still more at-issue than OVGs. In the second study, it was investigated whether CVGs are more salient than OVGs. The results tentatively suggest that there are no salience differences between CVGs and OVGs. Overall, the findings provide additional evidence in favor of a gradient approach to at-issueness.

Keywords: gesture semantics, at-issueness, perspective in gesture.

1. Viewpoint gestures

There is a long-standing tradition of investigating perspective-taking phenomena in speech. More recent research has shown that co-speech gestures can also encode perspective (McNeill, 1992). Verbal utterances are often accompanied by gestures, which can either be manual, i.e., performed with the hands and potentially other body parts, or facial, i.e., performed with the face. These gestures are often synchronized with the verbal expressions they co-occur with. The stroke (= the core) of a gesture, for example, is usually aligned with the nuclear accent of a word (Loehr, 2004; Ebert et al., 2011). Moreover, different alignment patterns of gesture and speech have been shown to have different semantic effects (Ebert and Ebert, 2014). This claim has been experimentally validated in a study reported in Ebert et al. (2022).

Previous research has distinguished different gesture types (for an overview, see McNeill, 1992), among them iconic gestures. Iconic gestures visually resemble a property of an object or action they illustrate. Perspective is often encoded in iconic gestures. McNeill (1992) distinguishes between CVGs, OVGs, and so-called *dual viewpoint gestures* (see also Parrill, 2010, 2012; Stec, 2012).

CVGs illustrate an event from a first-person perspective and the whole body is usually involved in the production of the gesture. OVGs, by contrast, illustrate an event as if observed from a distance and therefore only the hands and arms are involved when producing the gesture. For illustration, consider the cartoon scene in Figure 1 (taken from Parrill, 2010). In a study, Parrill (2010) let participants first watch different cartoon clips. Their task was to re-tell the cartoon scene to a friend they had brought to the study. This friend did not watch the cartoon scene before.

In Figure 1 one can see a skunk hopping across the room. If speakers re-tell this cartoon to

 In: Baumann, Geraldine, Daniel Gutzmann, Jonas Koopman, Kristina Liefke, Agata Renans, and Tatjana Scheffler (eds.) 2024. Proceedings of Sinn und Bedeutung 28. Bochum: Ruhr-University Bochum, 943-960.

Figure 1: Cartoon scene of a skunk hopping across a room. Taken from Parrill (2010)

their friend, they have several options to include gestures in their description, two examples are given in Figure 2. In Figure 2a, one can see a clear instance of a CVG because the speaker uses their whole (upper) body to depict the skunk's hopping movement from a first-person perspective. Figure 2b, by contrast, describes the same event but by means of an OVG. Here, the speaker only uses their index finger to trace the trajectory of the skunk. In other words, the speaker reports the hopping event from a third-person perspective. CVGs have been argued to be more informative than OVGs (Beattie and Shovelton, 2002). There is a third type of viewpoint gesture, which occurs very infrequently, however. This gesture type encodes multiple viewpoints at the same time and has therefore been dubbed dual viewpoint gesture (Parrill, 2009). Encoding multiple viewpoints in gesture seems to be less constrained than the occurrence of multiple viewpoints in speech since dual viewpoint gestures allow for the presence of two different character viewpoints at the same time, which has not been attested in spoken or written language. Interestingly, this is also possible in sign languages (e.g., Maier and Steinbach, 2022). Therefore, this might be a modality-specific feature. The co-presence of a character's and an observer's viewpoint in a gesture has been argued to be possible although it often produces an ironic effect because the two viewpoints seem to compete with each other (McNeill, 1992). Moreover, dual CVGs are restricted to specific contexts at least for adults (McNeill, 1992: p. 125). More specifically, one of the CVGs always is a deictic gesture to the speaker's body, representing the viewpoint of one character, and the body represents another character viewpoint. It can be noted, in sum, that the expression of multiple viewpoints is more liberal in gesture as opposed to speech.

2. At-issueness

2.1. Multiple meaning dimensions

Certain linguistic items and structures, such as *conventional implicatures* (CIs) and presuppositions, have long been recognized as contributing meaning in a different dimension than asserted content. In contrast to assertions, they have been considered to make not-at-issue contributions (cf. Koev, 2018). Not-at-issue content is overall considered to be a secondary meaning contribution whereas the asserted at-issue content makes the primary meaning contribution in an utterance, i.e., it is on the table for discussion (Farkas and Bruce, 2010).

The two types of meaning contributions differ in their discourse behavior. For example, one cannot directly assent or dissent with not-at-issue content, which is illustrated in (1) for appositives, which are not-at-issue by default. The asserted content, i.e., the proposition expressed by the main clause, by contrast, can be assented or dissented with directly (for further diagnostics testing for the at-issue/not-at-issue distinction see Tonhauser, 2012).

(1) A: My friend Sophie, a classical violinist, played a piece by Bach.
B: # No, she's not a classical violinist.
B': Hey, wait a minute! Sophie's not a classical violinist. (Koev, 2018)

(1) shows that directly denying the appositive's not-at-issue content (cf. the utterance of B) is not possible. Instead, a discourse-interrupting element such as *Hey, wait a minute!* has to be used to target the proposition expressed by the appositive (Shanon, 1976; von Fintel, 2004).

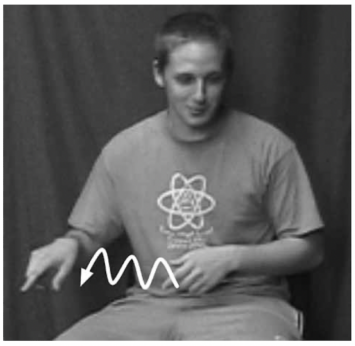
A further defining property of not-at-issue content is that it projects through negation and other entailment-canceling operators (Simons et al., 2010). (2) illustrates this.

(2) It is not the case that my friend Sophie, a classical violinist, played a piece by Mozart.
a. She performed a piece by Bach.
b. # She is a trumpet player.

When hearing an utterance as in (2), listeners have the intuition that the appositive's content escapes the scope of the negation, meaning that only the main clause proposition is targeted by the negation. This is illustrated by the continuations in (2a) and (2b), respectively. The continuation in (2a) picks up the at-issue main clause proposition, whereas the continuation in (2b) picks up the not-at-issue proposition expressed by the appositive. Since the main clause proposition is targeted by the sentential negation, (2a) is felicitous. (2b), by contrast, is infelicitous because it picks up the not-at-issue proposition expressed by the appositive, which projects through negation.

In the literature on at-issueness, at least three different notions of at-issueness can be distinguished (for an overview, see Koev, 2018). In this work, at-issueness will be defined by means of relevance to the *question under discussion* (QUD) (cf. Simons et al., 2010). A definition is given in (3).

(a) CVG used to depict the skunk in Figure 1 (Figure 3 in Parrill, 2010: p. 652)



(b) OVG used to depict the skunk in Figure 1 (Figure 2 in Parrill, 2010: p. 651)

Figure 2: Examples of a CVG and an OVG to depict the event shown in Figure 1

(3) a. A proposition *p* is at-issue iff the speaker intends to address the QUD via *?p*.
b. An intention to address the QUD via *?p* is felicitous only if:
(i) *?p* is relevant to the QUD, and
(ii) the speaker can reasonably expect the addressee to recognize this intention.
(Simons et al., 2010: p. 323)

Simons et al. (2010) thus define QUD relevance by means of a yes/no question that is associated with a proposition *p*, whereby the partition of the set of worlds into *p* and ¬*p* is denoted by *?p*. A question is defined as QUD-relevant iff it entails either a complete or a partial answer to the QUD. This means that *?p* must contain at least a partial answer to the QUD in order to be considered QUD-relevant. Part (ii) means that the at-issue status is determined by a further factor, namely the constraint that the speaker's intention to address the QUD is also recognized by the addressee. This part has to do with information packaging in general. Consider (4):

(4) Q: What is the weather like?
A: # Bob realizes that it's raining.
(Simons et al., 2010: p. 319)

The use of the factive verb *realize* in A's answer implies that the embedded proposition that it is raining is true, thus giving at least a partial answer to the QUD. However, this information is not asserted, but presupposed and thus not intended by the speaker to address the QUD, yielding the infelicity of A's response in (4). This example illustrates that the structural choice of the speaker also influences the at-issue status of a proposition. Projective content, such as the presupposition in (4), is predicted to be not-at-issue by Simons et al.'s (2010) definition.

Coming back to example (1), this means, then, that the proposition expressed by the main clause is intended to make an at-issue contribution by the speaker. The appositive's content, by contrast, is predicted to be not-at-issue. In other words, (1) can only be used felicitously if the QUD is such that the proposition expressed by the main clause (partially) answers it:

(5) Q1: What did your friend Sophie do?
Q2: What instrument does Sophie play?
A: My friend Sophie, a classical violinist, played a piece by Bach.

If A utters the sentence to address Q2, the whole utterance is infelicitous because the QUD targets the information provided by the appositive which are normally not used to address the QUD. However, if A intends to respond to a QUD along the lines of Q1, the utterance is felicitous because the information answering the QUD is provided by the main clause, i.e., the at-issue content.

2.2. The at-issue status of co-speech gestures

Ebert and Ebert (2014) analyze co-speech gestures as conventional implicatures and are thus predicted to behave similar to appositives (Potts, 2005). Therefore, they also make not-at-issue meaning contributions by default and share the projective behavior attested for appositives (cf. Section 2.1). Example (6) illustrates that co-speech gestures, just like appositives, cannot be directly denied in discourse.

(6) A: I brought [a bottle of water][1] to the talk. + "big" co-speech gesture
B: # No, the bottle isn't big.
B': Hey, wait a minute! Actually, the bottle isn't that big.
(cf. Ebert and Ebert, 2014)

Moreover, co-speech gestures also display the projective behavior attested for appositives:

(7) It is not the case that I brought [a bottle of water] to the talk. # A small one is enough for me.
+ "big" co-speech gesture

Here again the continuation is infelicitous as it targets the semantic content contributed by the gesture. Since this content is not-at-issue and therefore projective, the continuation is infelicitous.

The at-issue status of co-speech gestures has been empirically investigated by Ebert et al. (2020). In a rating study, participants had to rate how well a picture matched a videotaped utterance they had previously seen. In the utterance, either a co-speech gesture occurred giving information about an object's size or shape or this information was given by means of an adjective. An example is given in (8).

(8) a. In diesem Bild ist eine Mauer mit [einem Fenster] zu sehen. + "circular" co-speech gesture
b. In diesem Bild ist eine Mauer mit einem runden Fenster zu sehen.
'In this picture you can see a wall with a (round) window.'
(cf. Ebert et al., 2020: p. 173)

The picture participants were presented either matched or mismatched the information provided by the co-speech gesture or adjective. Thus, participants either saw a picture with a round window (matching) or a picture with a rectangular window (mismatching) for (8). Crucially, the match/mismatch only came about via the gesture or adjective. Assuming that not-at-issue content has a less severe impact on the overall truth conditions than at-issue content has (Kroll and Rysling, 2019), Ebert et al. (2020) hypothesized that the mismatch effect, i.e., the difference in mean ratings between the matching and mismatching condition, is higher for items in the adjective condition than for items in the gesture condition. Their results confirm their hypothesis, thus suggesting that co-speech gestures contribute not-at-issue meaning by default, in line with Ebert and Ebert (2014).

In a second experiment, Ebert et al. (2020) tested for the potential of demonstratives as so-called *dimension shifters*. This means that when a co-speech gesture is aligned with a demonstrative (such as German *so* 'such'), it shifts the gesture's semantic content from the not-at-issue to the at-issue dimension. In order to test for this, they compared the gesture and adjective condition from the first study to a third condition where the co-speech gesture was aligned with *so*. It was hypothesized that the ratings for the *so* + gesture condition were equal to the ratings of the (at-issue) adjective condition. The results, however, showed that the mismatch effect for the *so* + gesture condition was lower than for the adjective condition, but higher than for the gesture condition. This and similar findings for so-called *ideophones* (Barnes et al., 2022) eventually lead to the proposal to capture at-issueness as a gradient notion instead of a binary one (Barnes

[1] Square brackets indicate gesture-speech alignment.

and Ebert, 2023). Concretely, Barnes and Ebert (2023) propose that every expression comes with an inherent at-issue status. This can be shifted, however, if this expression addresses the QUD. In Section 5, their approach will be used to account for the results obtained from the experimental studies summarized in Sections 3 and 4 of the present paper.

It has been shown so far that utterances can contribute meaning in multiple dimensions: the asserted at-issue dimension and the backgrounded not-at-issue dimension. In general, projective content has been argued to contribute not-at-issue meaning by default (e.g., Simons et al., 2010). Ebert and Ebert (2014) have proposed to analyze co-speech gestures as conventional implicatures. Therefore, they are predicted to make not-at-issue contributions by default (for experimental validation see Ebert et al., 2020). The question that now arises is why CVGs and OVGs should differ in their at-issue status if they occur as co-speech gestures. A crucial difference between the gestures tested for in Ebert et al.'s (2020) studies is that they only used so-called no viewpoint gestures (Parrill, 2010), i.e., gestures that do not encode viewpoint at all. Moreover, as has been shown above, especially CVGs differ from the gestures used by Ebert et al. (2020) with respect to size as the whole body is involved in the production of a CVG (McNeill, 1992). In addition, it has been demonstrated that CVGs and OVGs also differ in certain ways: CVGs are more informative than OVGs (Beattie and Shovelton, 2002). Finally, in a study investigating the hypothesis that the perspective expressed in gesture and speech should be aligned by default, Hinterwimmer et al. (2021) found a CVG preference regardless of the perspective expressed in the speech signal. This preference might be due to a difference in the at-issue status of CVGs and OVGs. Based on these differences, it is hypothesized that although CVGs and OVGs both contribute not-at-issue meaning by default, the former type of viewpoint gesture is still more at-issue than the latter. This can be captured when assuming a gradient approach to at-issueness (Barnes and Ebert, 2023). Experiment 1 investigates this hypothesis.

3. Experiment 1: The at-issue status of viewpoint gestures

3.1. Method

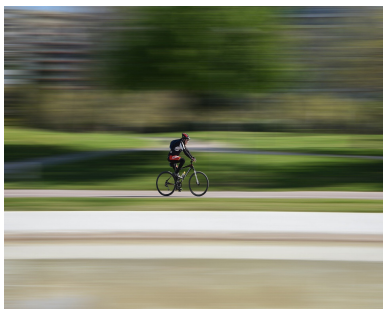
3.1.1. Participants

Participants (n = 60) were recruited via Prolific. All of them were self-reported native speakers of German and naive with respect to the research question.

3.1.2. Materials

Eighteen experimental items were created for the experiment. Each item consisted of two sentences: the first sentence introduced an event by the speaker and the second sentence further described this event. The second sentence contained a word, usually a verb, which was either accompanied by a CVG or an OVG. Alternatively, a phrase paraphrasing the gesture information preceded or followed this word (factor MODE: CVG vs. OVG vs. Verbal). The two gestures differed semantically only with regard to the perspectival information they encoded. An example for an experimental item can be found in (9).

(a) Matching picture



(b) Mismatching picture

Figure 3: The matching and the mismatching picture for the experimental item in ((9))

(9) a. Letzten Mittwoch hatte ich den ganzen Tag Termine überall in der Stadt. Nachdem einer der Termine länger dauerte als gedacht, musste ich mich [richtig beeilen.]
+ **CVG:** running using the whole body, or
+ **OVG:** running using two fingers

b. Letzten Mittwoch hatte ich den ganzen Tag Termine überall in der Stadt. Nachdem einer der Termine länger dauerte als gedacht, musste ich mich richtig beeilen **und rennen**.
'Last Wednesday I had many appointments throughout the whole city. After one of the appointments took longer than expected, I had to hurry a lot (and run).'

The part in bold print in (9b) indicates the paraphrase of the gestures' meaning in (9a).

Additionally, in order to establish either a match or a mismatch, each experimental item was paired with two pictures: a picture matching the utterance and a picture where there was a mismatch between utterance and picture (factor MATCH: match vs. mismatch). The matching picture used for the experimental item in (9) can be found in Figure 3a and the mismatching picture is given in Figure 3b. The study was thus of a 3x2 design which was similar to the design of the studies reported in Ebert et al. (2020).

The experimental items were interspersed with 24 unrelated filler items which were taken from Ebert et al. (2022). In the fillers, lexically ambiguous nouns were either paired with a gesture matching or mismatching the salient interpretation of the noun in the given sentence. These sentences were then also paired with a picture either matching or mismatching the salient interpretation of that noun.

3.1.3. Procedure

Participants had to complete a web-based questionnaire where they had to rate how well a videotaped utterance matched a picture they had seen prior to the video. The completion took approximately 20 minutes and participants were compensated with £3.20.

The questionnaire started with an introductory text describing the participants' task. They were instructed in this text to pay attention to the audio tape and the video tape when watching the videos. Additionally, the text informed them about their data protection and they were

also explicitly instructed that they can stop the questionnaire any time without providing any explanation.

The questionnaire was created using SoSci Survey (Leiner, 2022). Items were split up according to a Latin square design. Participants had to rate how well the utterance in the video matched the picture on a Likert scale ranging from 1 (= 'no match') to 7 (= 'perfect match'). They were asked in the beginning to judge each pair of video and picture individually and intuitively.

3.2. Predictions

Since CVGs differ more from the gestures used in Ebert et al. (2020) than OVGs do (cf. Section 1), first the negation test will be applied to an example with a CVG. The negation test is suitable to determine whether or not gestures or, more generally, parts of an utterance project and therefore contribute at-issue meaning.

(10) a. Simon overslept today. He then had to [hurry] to get to work on time.
+ CVG running (whole body)
b. #No, he did not run, he went by bicycle.
c. Hey, wait a minute! He had to hurry to get to work on time, but he did not run.

Example (10) shows that the denial of the content of the gesture in (10a) is not possible with the direct denial as in (10b). It is possible, however, to target the gesture when one uses the discoruse-interrupting *Hey, wait a minute!*, as in (10c). Therefore, (10) indicates that CVGs project, which has been argued to be behavior that is attested for any kind of not-at-issue content. It is thus hypothesized that CVGs, just as other kinds of speech-accompanying gestures, contribute meaning that is not-at-issue by default. This hypothesis is in line with the findings of the first experiment reported in Ebert et al. (2020). A significant interaction of the pairwise comparison of CVG and Verbal for the factor MODE and the factor MATCH is predicted. In other words, this means that the rating difference between matching and mismatching CVG items is predicted to be significantly lower than the rating difference between matching and mismatching Verbal items. In the terms of Ebert et al. (2020), this means that the mismatch effect is predicted to be stronger for Verbal items than for CVG items.

In order to determine whether there CVGs and OVGs differ with respect to their at-issue status, consider (11), a slightly modified version of (10).

(11) a. Simon overslept today. He then had to [hurry] to get to work on time.
+ **CVG:** running (whole body), or
+ **OVG:** running (with two fingers)
b. #No, he did not run, he went by bicycle.
c. Hey, wait a minute! He had to hurry to get to work on time, but he did not run.

Example (11) shows that direct denial is possible neither for OVGs nor for CVGs. Therefore, the comparison suggests that CVGs and OVGs behave alike concerning their at-issue status. However, due to a general preference of CVGs found by Hinterwimmer et al. (2021), the higher informativity of CVGs (Beattie and Shovelton, 2002), and possibly also due to the

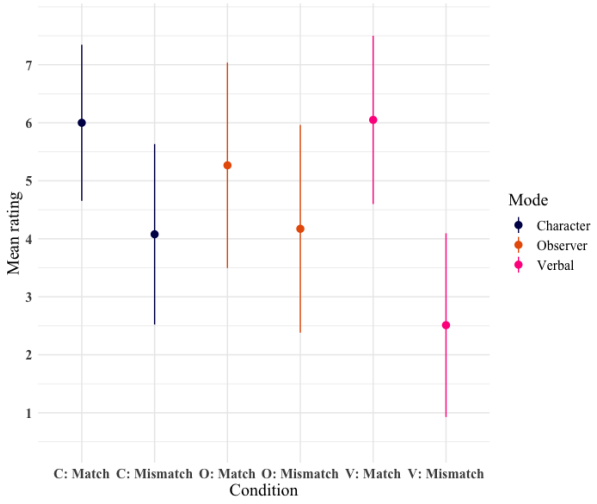


Figure 4: Mean values and SDs for each condition

larger size of CVGs (McNeill, 1992) a difference in the at-issue status of CVGs and OVGs is hypothesized. This hypothesis leads to a problem. Traditionally, as shown above, the notion of at-issueness has been understood as a binary notion, meaning that an expression or a gesture is either at-issue or not-at-issue (e.g., Simons et al., 2010). If it were assumed that at-issueness is a binary notion, this would lead to the prediction that CVGs are at-issue. However, as has been shown above already, this is not a desirable prediction. Therefore, it will be assumed that at-issuness is a gradient notion (Barnes and Ebert, 2023). CVGs and OVGs can thus differ in their at-issue status and still both contribute not-at-issue meaning by default at the same time. A significant interaction of MATCH and the pairwise comparison between CVG and OVG of the factor MODE is thus predicted. This means that the rating difference between matching and mismatching items is predicted to be significantly higher for CVG items than for OVG items, i.e., the mismatch effect is predicted to be stronger for CVG items than for OVG items.

3.3. Results

The statistical analysis was conducted using RStudio (RStudio Posit Team, 2023), an integrated development enviroment for the R statistics software (R Core Team, 2022). For data preparation and visualization, the package 'tidyverse' (Wickham et al., 2019) was used. In order to test for significant effects, the results were analyzed using cumulative link mixed effects models with the function clmm() in the 'ordinal' package (Christensen, 2019). An ordinal mixed effects model was chosen to analyze the data instead of a linear mixed effects model mainly for two reasons: i) linear mixed effects models require the data to be measured at the interval level and ii) linear mixed effects models require the data to be normally distributed. Both is questionable for Likert scale data. Therefore, ordinal mixed effects models are more suitable to analyze this kind of data. For the statistic computation, the two factors and all the interactions between them were entered as fixed effects into the model using effect coding, i.e., the intercept represents the unweighted grand mean while the fixed effects compare the factor levels to each other. For MODE, the three-level factor, two pairwise contrasts were defined. Pairwise contrasts compare two levels of a factor with each other.

Table 1: Ordinal mixed effects model with Mode and Match as fixed effects and participants and items as random intercepts. Formula: 'responseO ∼ Match * Mode + (1 | CASE) + (1 | nr)', data = experimental, Hess = T, nAGQ = 1
Signif. codes: *** 0.001 | ** 0.01 | * 0.05 | . 0.1

	Estimate	Std. Error	z value	Pr(> |z|)
Match	-2.857	0.139	-20.60	<2e-16 ***
Mode – CVG vs. OVG	-0.423	0.136	-3.11	0.0019 **
Mode – OVG vs. V	0.394	0.143	2.76	0.0057 **
Match:Mode – CVG vs. OVG	1.176	0.273	4.30	1.7e-05 ***
Match:Mode – OVG vs. V	3.193	0.293	10.90	<2e-16 ***

Table 2: Ordinal mixed effects model with Mode and Match as fixed effects and participants and items as random intercepts. Formula: 'responseO ∼ Match * Mode + (1 | CASE) + (1 | nr)', data = experimental, Hess = T, nAGQ = 1
Signif. codes: *** 0.001 | ** 0.01 | * 0.05 | . 0.1

	Estimate	Std. Error	z value	Pr(> |z|)
Match	-2.857	0.139	-20.60	<2e-16 ***
Mode – CVG vs. OVG	-0.423	0.136	-3.11	0.0019 **
Mode – CVG vs. V	-0.816	0.144	-5.65	1.6e-08 ***
Match:Mode – CVG vs. OVG	1.176	0.273	4.30	1.7e-05 ***
Match:Mode – CVG vs. V	-2.018	0.289	-6.99	2.7e-12 ***

Figure 4 shows the mean ratings and the standard deviations (SDs) for each condition. CVG items were rated similar to Verbal items in the match condition (CVG: M = 6.00, SD = 1.35; Verbal: M = 6.05, SD = 1.45). OVG items were rated worse than the other two conditions of the factor MODE in the match condition (M = 5.27, SD = 1.77). In the mismatch condition, CVG and OVG items were rated similarly (CVG: M = 4.08, SD = 1.56; OVG: M = 4.17, SD = 1.79). Verbal items were rated worse in this condition (M = 2.51, SD = 1.58).

The mixed effects model corresponding to the data in Figure 4 is given in Table 1. There is a main effect for the factor MATCH. For the pairwise comparison between CVG and OVG for the factor MODE, there is also a main effect. Moreover, there is a main effect for the pairwise comparison between OVG and Verbal (abbreviated as V in Tables 1 and 2) for the factor MODE. In addition, a significant interaction between MATCH and the pairwise comparison between CVG and OVG for MODE can be observed. Finally, there is an interaction for the factor MATCH and the pairwise comparison for MODE between OVG and Verbal.

For safety reasons, although uncommon, the same model was computed with different pairwise contrasts. While the pairwise contrasts for the factor MODE in Table 1 compared CVG to OVG and OVG to Verbal, the pairwise contrasts for the model presented in Table 2 compared CVG to OVG and CVG to Verbal. Table 2 shows that there is a main effect for the factor MODE for the pairwise comparison between CVG and Verbal. Moreover, a significant interaction between the factor MATCH and the pairwise comparison between CVG and Verbal for the factor MODE can be observed.

3.4. Discussion

The results show a significant rating difference between CVG and Verbal items. Moreover, a significant rating difference in the rating of matching vs. mismatching CVG items and Verbal items can be observed. This is the interaction predicted for Research Question 1, which targets the at-issue status of CVGs. Therefore, the data confirm the hypothesis that CVGs contribute not-at-issue meaning by default. This finding is also in line with the results of the study reported in Ebert et al. (2020) suggesting that co-speech gestures contribute not-at-issue meaning by default.

Concerning the hypothesis that CVGs and OVGs differ with respect to their at-issue status, a significant interaction between MATCH and the pairwise comparison between CVG and OVG for the factor MODE was predicted. Again, the results of the study confirm this prediction. The rating difference between matching and mismatching items in the CVG condition was significantly higher than in the OVG condition. Therefore, it can be concluded that CVGs are more at-issue than OVGs. This interpretation is also in line with gradient approaches to at-issueness (cf. Barnes and Ebert, 2023).

A potential confounding factor is that there might be salience differences between CVGs and OVGs as the former are normally larger in size due to the whole body being involved in the gesture production (McNeill, 1992). Therefore, a second experiment was conducted targeting the research question whether the two viewpoint gesture types differ with respect to salience.

4. Experiment 2: Are there salience differences between viewpoint gestures?

4.1. Salience: An overview

Very broadly, salience can be described as one of the main organizing principles in discourse (Falk, 2014). This means that during discourse processing some parts of an utterance are more activated during discourse processing. These more activated parts are seen as salient units in a discourse.

One can distinguish two notions of salience in linguistic research: *backward-looking* and *forward-looking* salience. The underlying assumption for backward-looking salience is that salient units can be retrieved from memory more easily than discourse units which are not salient. Research on backward-looking salience has in large parts focused on referring expressions. It has been proposed that these are ranked on a hierarchical list, whereby each of them has an accessibility value (Ariel, 1990). Speakers choose referring expressions to refer back to a discourse unit in accordance with that accessibility value:

(12) Adele went to the university library.
  a. She urgently needed a book for her exams.
  b. This institution was one of the oldest in the country.

(Falk, 2014: p. 3)

Adele is highly salient in (12). Therefore, this proper name is picked up with a personal pronoun in (12a) because personal pronouns are used to pick up discourse units which are highly

accessible. In order to pick up *the university library*, a discourse unit which is not salient, by contrast, speakers are predicted to use a demonstrative DP (*this institution* in (12b)) instead of a personal pronoun.

By contrast, forward-looking salience relates to the idea that salience is associated with the allocation of attention (e.g., Grosz et al., 1995). Salience under this notion is seen as means of attentional control (Chiarcos, 2009). This can be illustrated with the so-called *Moses-illusion*:

(13) Moses took two animals of each species to the Ark.

The sentence in (13) is false, since Noah took two animals of each species to the Ark. However, when asked to judge this sentence as true or false, speakers often judge it true. This is because according to the notion of forward-looking salience *Moses* is backgrounded in this example. If, by contrast, speakers are presented with the slightly modified example in (14), they detect the error more frequently and judge the sentence false (Bredart and Modolo, 1988).

(14) It was Moses who took two animals of each species to the Ark.

Thus, foregrounding a discourse unit can affect its salience.

The question that now arises is how these ideas about salience can be applied to the present case of viewpoint gestures. CVGs could in principle be more salient than OVGs at least on the perceptual level because they are larger in size. However, it has been shown in Section 2.2 that co-speech gestures contribute not-at-issue meaning by default. This has been verified for CVGs and OVGs by the results of Experiment 1 (Section 3). Not-at-issue content is generally backgrounded material as it is not asserted (cf. Potts, 2005). Moreover, the manual modality is the backgrounded modality in spoken languages. Therefore, fit is plausible to assume under the forward-looking account of salience that CVGs and OVGs both are non-salient units in a discourse. While the predictions of the backward-looking notion of salience are somewhat less clear, CVGs should be more easily retrievable from memory if they are more salient than OVGs. This is tested for by means of a forced choice study. Due to the high backgroundedness of both CVGs and OVGs, it is hypothesized that they do not differ with respect to their salience, i.e., that they are both not salient.

4.2. Method

4.2.1. Participants

Self-reported native speakers of German were recruited as participants (n = 45) via Prolific. All of them were naive with respect to the research question.

4.2.2. Materials

The same materials as for Experiment 1 were used, i.e., participants were presented videotaped utterances (n = 18) in the style of (9) that were either accompanied by a CVG or an OVG (cf. (9a)) or a verbalized version of the gestures as in (9b) (factor MODE: CVG vs. OVG vs.

Verbal). These utterances were paired with a matching and a mismatching picture. Crucially, as in Experiment 1, the (mis)match came about only via the manipulation of the factor MODE. The experimental items were interspersed with 22 unrelated fillers. In order to increase the overall difficulty of the study and thus to avoid a ceiling effect, a distractor task was included where participants had to memorize a randomly generated digit string consisting of six digits. The questionnaire was thus of a single factor design.

4.2.3. Procedure

Participants had to complete a web-based forced choice questionnaire. There task was to select the picture matching the utterance they were previously presented. The completion took approximately 20 minutes and participants were compensated with £3.20.

The questionnaire started with an introduction where participants were instructed with the task of the study. They were specifically instructed to pay attention to both the audio and the video tape. In addition, they were informed about their data protection rights and had to give informed consent prior to participation.

Each trial started with the presentation of the randomly generated six digit string. Participants saw the digit string for five seconds. Afterwards, the questionnaire automatically proceeded to the next page where the video played after three seconds. The video could only be viewed once. Then, on the next page, participants were presented three digit strings. Their task was to select the string they had seen at the beginning of the trial. On the last page of each trial, two pictures were presented. One picture matched the utterance they were previously presented and one did not match the utterance. The task of the participants here was to select the matching picture.

4.3. Predictions

As has been shown in Section 2.2, co-speech gestures contribute not-at-issue meaning by default (cf. Ebert et al., 2020). Not-at-issue content has been argued to be backgrounded information (Potts, 2005). Moreover, salience is associated with allocation of attention (e.g., Grosz and Sidner, 1986; Grosz et al., 1995) and backgrounded information is never salient. Thus, it follows that not-at-issue content is not salient. Since co-speech gestures are not only not-at-issue but also occur in the visual modality, the backgrounded modality in spoken languages, it is assumed that this makes them even more backgrounded. It is therefore hypothesized that CVGs and OVGs do not differ with regard to their salience as they both are highly backgrounded in a multimodal speech signal. TA main effect for the factor MODE is thus predicted,, i.e., that the proportion of false choices is significantly lower in the condition Verbal compared to the two gesture conditions because spoken material is always more foregrounded than gestural material.

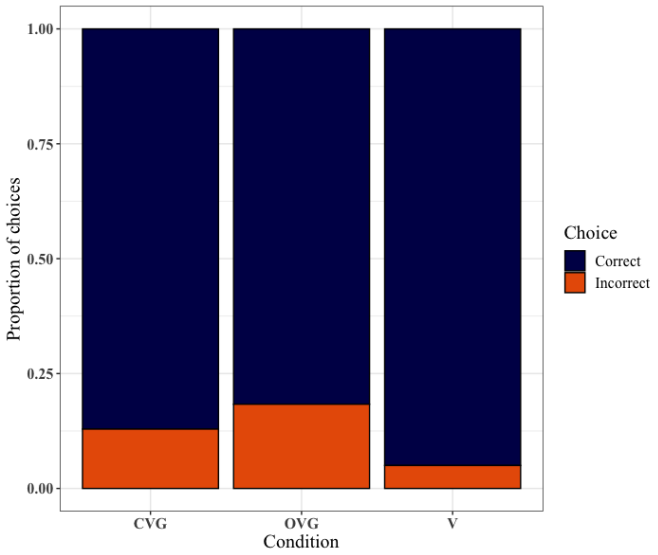


Figure 5: Proportions of choices (correct/incorrect) for each condition

Table 3: Logistic regression model with Mode as fixed effect and participants and items as random intercepts. Formula: mismatch ∼ Mode + (1 | CASE) + (1 | item), control = glmerControl(optimizer = "bobyqa"), family = binomial(link = "logit"), wo_items
Signif. codes: *** 0.001 | ** 0.01 | * 0.05 | . 0.1

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	-2.415	0.246	-9.84	<2e-16 ***
Mode – OVG vs. CVG	0.468	0.27	1.73	0.0831 .
Mode – CVG vs. Verbal	1.126	0.369	3.05	0.0023 **

4.4. Results

The same software as for Experiment 1 was used for statistical analysis, data preparation, and visualization was used. In order to test for significance, the results were analyzed using a logistic regression model with the glmer() function in the package 'lme4' (Bates et al., 2015). For the statistic computation, the factor was entered as fixed effect into the model using contrast coding where two pairwise comparisons were defined, one comparing the OVG with the CVG condition and one comparing the CVG condition with the Verbal condition.

Two items were included from the statistical analysis since participants chose the wrong picture in the OVG condition in the majority of cases, thus suggesting that the gesture was interpreted falsely. The proportions of choices are given in Figure 5. Participants chose the wrong picture most often in the OVG condition (18.3%), followed by the CVG condition (12.9%), and least often in the Verbal condition (5%). The overall accuracy for the distractor task was 95.3%.

The model corresponding to the data in Figure 5 is given in Table 3. The model output shows that there is a significant difference in the pairwise comparison of the CVG and the Verbal condition. In addition, there is a marginally significant difference between the OVG and the CVG condition.

4.5. Discussion

The results show that, in line with the predictions, participants chose the wrong picture least often in the condition Verbal, the condition which was hypothesized to contain the most salient material. Descriptively, the results also suggest a difference between the CVG and OVG condition. However, this difference only reached marginal significance, thus tentatively suggesting that there are no relevant salience differences between the two gesture types, as was hypothesized. Future research is needed, however, to fully clarify this matter.

5. General discussion and conclusion

The results of two experimental rating studies were reported investigating i) the at-issue status of CVGs and OVGs when they occur as co-speech gestures (Experiment 1) and ii) potential salience differences between the two types of viewpoint gestures (Experiment 2). For Experiment 1, it was hypothesized that although CVGs and OVGs contribute not-at-issue meaning by default, they still differ with respect to their at-issue status. More concretely, it was hypothesized that CVGs are more at-issue than OVGs due to the higher informativity of the former gesture type (Beattie and Shovelton, 2002) and potentially also due to their larger size (McNeill, 1992). The results obtained from Experiment 1 (Section 3) lend support to this hypothesis as the mismatch effect was higher for items in the CVG than in the OVG condition. Moreover, the mismatch effect was highest for items in the condition Verbal, thus suggesting the default not-at-issue status of CVGs and OVGs.

A potential confounding factor in the aforementioned interpretation of the results of Experiment 1 is that the two types of viewpoint gestures differ with respect to their salience. Therefore, Experiment 2 was conducted. Since co-speech gestures are in general highly backgrounded not only due to their not-at-issue status but also due to the manual modality being generally the backgrounded one in spoken languages, it was hypothesized that there are no salience differences between the two gesture types. The results of Experiment 2 generally support the hypothesis that co-speech gestures are less salient than verbal material. However, the pairwise comparison between CVG and OVG items yields a marginally significant effect. Although this tentatively suggests that there are no salience differences between the two types of viewpoint gestures, further research is needed to fully clarify this.

Taken together, the results of the two experiments reported in this paper are interpreted as follows: since there seem to be no salience differences between CVGs and OVGs, they arguably differ in their at-issue status. The results of Experiment 1 generally lend support not only to Ebert and Ebert's (2014) theory but also to the experimental paradigm used by Ebert et al. (2020) since the design employed for Experiment 1 in this paper was a slightly adapted version of their design of Experiment 1.

From a theoretical perspective, the results of the studies reported here also substantiate the proposal that at-issueness is a gradient notion and not a binary one (Barnes and Ebert, 2023). In their proposal, Barnes and Ebert (2023) define at-issueness as being related to the QUD. However, they depart from standard QUD-based approaches to at-issueness (e.g., Simons et al., 2010) in that they assume that the at-issue status of a linguistic structure is not fixed but instead

can be shifted on a scale if there is some contextual pressure, i.e., if it addresses the QUD (for potential drawbacks of their approach see Koev, 2023). For speech-accompanying gestures and ideophones, they propose the following scale (p. 212),

(15) Iconic co-speech gestures > sentence-medial adverbial ideophones > DEM + iconic co-speech gestures > DEM + sentence-medial adverbial ideophones

where > reads as 'is less at-issue than'. The proposal put forth in this paper is to split up the first category, iconic co-speech gestures, into several subcategories where OVGs rank lower on the scale than CVGs. From this proposal it follows that when a CVG is aligned with a demonstrative, such as German *so* (cf. Section 2.2 of the present paper), it will shift more toward the at-issue dimension than when an OVG is aligned with a demonstrative. Therefore, CVGs are argued to be inherently more at-issue than OVGs, i.e., the higher mismatch effect for CVGs observed in Experiment 1 is not due to contextual factors. It is left to future research to determine which factors cause the higher degree of at-issueness of CVGs compared to OVGs.

It is for now unclear whether the assumption actually holds that the differences in the at-issue status of CVGs and OVGs are inherent. This unclarity stems from the design of Experiment 1 itself and the fact that the gradient approach to at-issueness proposed by Barnes and Ebert (2023) is a QUD-based one. In Experiment 1, participants were always presented the following question: *Wie passend finden Sie die Äußerung in Bezug auf das Bild?* ('How appropriate do you find the statement in relation to the image?'). Since the match or mismatch always came about only via the manipulation of the MODE factor, this question might have been interpreted as a QUD explicitly targeting the content of the gestures used in the experiment. An alternative interpretation of the results, then, is that the differences in the at-issue status of CVGs and OVGs are not inherent. Rather, both gesture types could be inherently equally not-at-issue, but CVGs can be shifted more toward the at-issue dimension than OVGs. A follow-up study with pending results investigating this alternative interpretation is currently being conducted.

References

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London, UK: Routledge.

Barnes, K. and C. Ebert (2023). The information status of iconic enrichments: Modelling gradient at-issueness. *Theoretical Linguistics 49*(3-4), 167–223.

Barnes, K., C. Ebert, R. Hörnig, and T. Stender (2022). The at-issue status of ideophones in German: An experimental approach. *Glossa: A Journal of General Linguistics 7*(1), 1–39.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67*(1), 1–48.

Beattie, G. and H. Shovelton (2002). An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology 93*(2), 179–192.

Bredart, S. and K. Modolo (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica 67*(2), 135–144.

Chiarcos, C. (2009). *Mental Salience and Grammatical Form: Toward a Framework for Salience Metrics in Natural Language Generation*. Ph. D. thesis, University of Potsdam, Potsdam, Germany.

Christensen, R. H. B. (2019). Ordinal—Regression Models for Ordinal Data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Ebert, C. and C. Ebert (2014). Gestures, demonstratives, and the attributive/referential distinction. Talk given at *Semantics and Philosophy in Europe 7*.

Ebert, C., C. Ebert, and R. Hörnig (2020). Demonstratives as dimension shifters. In M. Franke, N. Kompa, M. Liu, J. L. Mueller, and J. Schwab (Eds.), *Proceedings of Sinn und Bedeutung 24*, Osnabrück, pp. 161–178. University of Osnabrück.

Ebert, C., S. Evert, and K. Wilmes (2011). Focus marking via gestures. In I. Reich, E. Horch, and D. Pauly (Eds.), *Proceedings of Sinn und Bedeutung 15*, Saarbrücken, Germany, pp. 193–208. University of Saarland.

Ebert, C., G. Pirillo, and S. Walter (2022). The role of gesture-speech alignment for gesture interpretation. In S. Featherston, R. Hörnig, A. Konietzko, and S. von Wietersheim (Eds.), *Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data*, Tübingen, Germany, pp. 65–77. University of Tübingen.

Falk, S. (2014). On the notion of salience in spoken discourse-prominence cues shaping discourse structure and comprehension. *TIPA. Travaux interdisciplinaires sur la parole et le langage 30*, 1–23.

Farkas, D. F. and K. B. Bruce (2010). On reacting to assertions and polar questions. *Journal of Semantics 27*(1), 81–118.

Grosz, B. J., A. K. Joshi, and S. Weinstein (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics 21*(2), 203–225.

Grosz, B. J. and C. L. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics 12*(3), 175–204.

Hinterwimmer, S., U. Patil, and C. Ebert (2021). On the interaction of gestural and linguistic perspective taking. *Frontiers in Communication 6*, 1–15.

Koev, T. (2018). Notions of at-issueness. *Language and Linguistics Compass 12*(12), e12306.

Koev, T. (2023). Gradient at-issueness versus uncertainty about binary at-issueness. *Theoretical Linguistics 49*(3-4), 249–260.

Kroll, M. and A. Rysling (2019). The search for truth: Appositives weigh in. In K. Blake, F. Davis, K. Lamp, and J. Rhyne (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 29*, Los Angeles, CA, pp. 180–200. University of California, Los Angeles.

Leiner, D. J. (2022). SoSci Survey (Version 3.3.14a). Available at https://www.soscisurvey.de.

Loehr, D. P. (2004). *Gesture and Intonation*. Ph. D. thesis, Georgetown University, Washington, DC.

Maier, E. and M. Steinbach (2022). Perspective shift across modalities. *Annual Review of Linguistics 8*(1), 59–76.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.

Parrill, F. (2009). Dual viewpoint gestures. *Gesture 9*(3), 271–289.

Parrill, F. (2010). Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes 25*(5), 650–668.

Parrill, F. (2012). Interactions between discourse status and viewpoint in co-speech gesture. In B. Dancygier and E. Sweetser (Eds.), *Viewpoint in Language: A Multimodal Perspective*, pp. 97–112. Cambridge, UK: Cambirdge University Press.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford, UK: Oxford University

Press.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

RStudio Posit Team (2023). *RStudio: Integrated Development Environment for R*. Boston, MA: Posit Software, PBC.

Shanon, B. (1976). On the two kinds of presuppositions in natural language. *Foundations of Language 14*(2), 247–249.

Simons, M., J. Tonhauser, D. Beaver, and C. Roberts (2010). What projects and why. In N. Li and D. Lutz (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 20*, Ithaca, NY, pp. 309–327. CLC Publications.

Stec, K. (2012). Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture 12*(3), 327–360.

Tonhauser, J. (2012). Diagnosing (not-)at-issue content. In E. Bogal-Allbritten (Ed.), *Proceedings of Semantics of Under-represented Languages of the Americas (SULA) 6*, pp. 239–254.

Tonhauser, J., D. I. Beaver, and J. Degen (2018). How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics 35*(3), 495–542.

von Fintel, K. (2004). Would you believe it? The king of France is back! In A. Bezuidenhout and M. Reimer (Eds.), *Descriptions and Beyond*, pp. 315–347. Oxford, UK: Oxford University Press.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software 4*(43), 1686.