

And the Oscar for the best supporting actor goes to...*

Cornelia Ebert, Universität Frankfurt

“Working as a supporting artiste is not as glamorous as it sounds. [...] Extras aren’t supposed to say anything during a take, for as the saying goes, you aren’t being paid to talk.”¹

In this short paper, I will be concerned with what I consider the supporting actors in language: intonation and gesture. I will present the results of an experiment that aims at contributing to the question what supporting role they play exactly and finally settling the question which of the two will be awarded the Oscar.

The idea for this contribution grew out of a conversation I had with Caroline some weeks ago, where she asked me about my motivation to engage with gestures, since – so she conjectured (in line with many others I had talked to about this before) – gestures cannot belong to language proper, one argument being that people happily communicate without relying on gestures, for example on the phone. I told her that I was surprised to hear this concern from her as a phonologist, because while it’s true that communication is possible without gesture it is also true that communication is possible without intonation, evidenced by the fact that we can understand written texts.

This contribution aims at following up on the conversation we had on that occasion, but also continuing earlier discussions when Caroline was the spokesperson of the SFB “Information Structure” in Potsdam (2003-2010), where I had a PhD position and had worked on semantic effects of information structure. In fact, I believe that gesture nowadays shares the same fate as intonation some 40 years ago, when the role of intonation and its semantic potential was still unsettled.

Theory and background. It took a while until generativist linguists acknowledged that intonation belongs to language and could possibly be an interesting field for investigation also in a formal realm. It took even longer until semanticists understood that intonation does not only affect appropriateness conditions, but also truth conditions and could hence be a fruitful topic of semantics and not only pragmatics. Rooth (1985) (but see already Dretske 1975) pointed out that (1a) has different truth conditions from (1b), while the only difference between the two sentences lies in their focus assignment, which is reflected by the fact that (1a) has its nuclear accent on BILL, while in (1b) it is on SUE.

- (1) a. *Mary only introduced [BILL]_F to Sue.*
b. *Mary only introduced Bill to [SUE]_F.*

(1a) is true in situations where there was no other person apart from Bill that was introduced to Sue, whereas (1b) is only true in situations in which Bill was introduced exclusively to Sue, but to nobody else. In other words, (1a) can be falsified by a situation where Mary introduced Bill and Paul to Sue, while (1b) would be true in such a situation, as long as Bill was not introduced to anybody else than Sue.

Ever since Rooth’ (1985) seminal thesis, semanticists investigated the different truth-conditional effects that information-structural categories like focus, but also topicality can have: Information structure has the power to determine a quantifier’s arguments, in particular adverbial quantifiers tend to interpret topical material in the restrictor and focal material in the

* This work has been conducted in the context of the project *PSIMS* (‘The Pragmatic Status of Iconic Meaning in Spoken Communication: Gestures, Ideophones, Prosodic Modulations’) within the Priority Program 1727 *XPRAG.de*, funded by the German Science Foundation DFG. For helpful discussions and comments as well as support with the experimental design, execution, and statistical evaluation I would like to thank Jörg Dreyer, Christian Ebert, Stefan Hinterwimmer, Robin Hörnig, Robert Müller, Giovanna Pirillo, Peter Staudacher, and Theresa Stender.

¹ <https://www.theguardian.com/commentisfree/2016/dec/26/secret-life-film-television-extra-supporting-artiste>

nuclear scope (Rooth 1985, Partee 1991, Herburger 1993, Krifka 1990, Diesing 1992). There is also a close connection between specificity, partitivity, and wide scope on the one hand and topicality on the other. Topical indefinites tend to be interpreted specifically or take wide scope (Cresti 1995, Jäger 1999, Endriss & Hinterwimmer 2006, Endriss 2009). And quite generally, information structure and accenting have a great influence on the availability of scope readings (Pafel 1997, Büring 1997).

Some decades later the same picture arises, I believe, in a different context. In these days it is common linguistic wisdom that intonation belongs to language proper, that it can be studied by using the same formal methods as other phenomena, and that it can have hard semantic truth conditional effects. But all this does not seem to hold for co-speech gestures. Among linguists, gestures are still widely seen as speech-accompanying material that is not linguistic in nature and that cannot be investigated and captured with the established formal linguistic methods. There seems to exist some silent agreement that while gesture might add some kind of meaning to what is said, this meaning component is necessarily different in nature from meaning that comes from speech and should not be investigated with the same tools. Furthermore, people often argue that gesture contributions cannot be of any significance because we can do without. We can communicate and understand each other just by talking without taking gestures into account, as we do on the phone.

But isn't this really the same situation as with intonation? When we read a text, understanding language is possible without taking intonation into account. And yet, we know that intonation is an integral part of language – and one, as we have pointed out above, that can make truth-conditional contributions. When we read, we consider language in a different modality, abstracted from its intonational properties. Why not consider spoken language without gesture in the same way – as language abstracted from its visual semantic potential?

In recent years, Lascarides & Stone (2009), Ebert & Ebert (2014) and Schlenker (2018), among others, have argued that co-speech gestures do contribute meaning in the standard semantic sense. And it could be shown that this gestural meaning enters composition as non-at-issue information by default. Consider an utterance of (2) (along the lines of examples from Ebert & Ebert 2014) with a simultaneous iconic gesture indicating an oval object (where underlining indicates co-occurrence of speech and gesture).

(2) *Caroline bought a casserole*_{+ 'oval' gesture}.

The verbal and gestural meaning together communicate that Caroline bought a casserole and that this casserole was oval. The gestural meaning comes in as non-at-issue meaning, evidenced by the fact that gestural meaning cannot be directly denied (cf. 3a) and that it projects across operators as negation, such that (3b) is an incoherent piece of discourse in contrast to (3c), where the shape information is expressed by an adjective.

- (3) a. #*That's not true, the casserole isn't oval./Hey, wait a minute, the casserole isn't oval.*
b. *I would never buy a casserole*_{+ 'oval' gesture}. #*Lasagna noodles do not fit in there.*
c. *I would never buy an oval casserole. Lasagna noodles do not fit in there.*

Although people have argued that non-at-issue material does not contribute to truth conditions in the same way as at-issue material (Potts 2005²), it is without any doubt that non-at-issue material has some influence on the truth conditions of a sentence containing it (but see Syrett & Koev 2014 for discussion). This also holds for the gesture contribution in (2). While a situation where Caroline bought an oval casserole verifies the utterance, one where she bought

² Ebert et al. (2019) investigate the nature of the semantic contribution of gestures and their influence on truth value judgements. They find experimental verification of their non-at-issueness in a picture verification task.

a square one does not in the same way, because it does not meet the requirements that come in via the gestural component, i.e. that the casserole Caroline bought was oval.

So under the hypothesis that gesture makes similarly meaningful contributions as intonation the question arises how strong the semantic impact of gesture and speech is and whether gesture can influence truth conditions in similar ways as intonation. And a further question would be: which of the two is the one that defeats the other and has the stronger semantic potential? We aimed at pursuing precisely this question by way of a forced choice study that investigates reference resolution of pronouns on the basis of co-speech gesture information on the one hand and intonation on the other.

From the literature we know that intonation can control pronoun resolution, which is yet another case of intonation influencing truth conditions. Below, I will show that the same holds for co-speech gesture, too. Lakoff (1971) was among the first to point out that intonation plays a crucial role for pronoun resolution, as evidenced by the following example.

- (4) a. *Paul insulted Bill. And then he hit him.*
b. *Paul insulted Bill. And then HE hit HIM.*

While the second sentence of (4a) is most naturally interpreted with the first pronoun (*he*) resolved to Paul and the second (*him*) to Bill, i.e. *parallel* to the presentation of the DPs in the first sentence, it is the other way around in (4b), where the pronouns carry contrastive stress. Here, pronoun resolution is *inverse* and *HE* is interpreted as referring to Bill and *HIM* to Paul.

Interestingly, pronoun resolution can also be managed by co-speech gestures. A speaker can point to a location in the gesture space while using a name to locate a referent in the gesture space (see Kendon 2004, Fricke 2009, cf. Bühler's *deixis am phantasma*) and to refer back to the respective referents via a pronoun (see Goodrich Smith & Hudson Kam 2012 and Foraker 2014 for experimental studies and Ebert 2018 for discussion of related cases). In the following examples, we consider a situation where the referents under discussion (Ede and Jan) are not present and hence a proper deictic use of the pronoun *he* is excluded. Here and in subsequent examples, A and B denote locations in the gesture space. A is a point to the left of the central gesture space in front of the speaker and B is a point more to the right of the central gesture space.

- (5) a. *Jan_{+point to A} talked to Ede_{+point to B} about emptying the dish washer and collecting premiums. And then he_{+point to A} left the room.*
b. *Jan_{+point to A} talked to Ede_{+point to B} about emptying the dish washer and collecting premiums. And then he_{+point to B} left the room.*

In (5a), *he* in the second sentence is most naturally resolved to Jan, while in (5b) it is resolved to Ede due to the co-speech gesture, which serves to determine reference and thus has truth-conditional effect in these cases.

Experiment: design. In an experimental study, we now aimed at testing for the reference determination potential of intonation, which we know from the literature, in comparison to the much less understood reference determination potential of co-speech gestures.

The study was designed as a forced choice experiment with four conditions in a Latin Square design, i.e. participants saw one item only in one of the four conditions. We had eight items (and nine filler items). Each participant hence saw two items per condition. We tested 16 German native speakers, 4 per list. Participants were recruited from the experiment platform *prolific* (<https://prolific.ac>).

Consider the following example item.

- (6) *Erst hat Peter Paul geÄRgert. Und dann hat er ihn geSCHLAgen.*

,First Peter annoyed Paul. And then he hit him.'

Such an item came in the following four conditions:

1. unstressed + no gesture

Erst hat Peter Paul geÄRgert. Und dann hat er ihn geSCHLagen.

2. stressed + no gesture

Erst hat Peter Paul geÄRgert. Und dann hat ER IHN geSCHLagen.

3. unstressed + gestures (inverse)

Erst hat Peter_{+point to A} Paul_{+point to B} geÄRgert. Und dann hat er_{+point to B} ihn_{+point to A} geSCHLagen.

4. stressed + gestures (parallel)

Erst hat Peter_{+point to A} Paul_{+point to B} geÄRgert. Und dann hat ER_{+point to A} IHN_{+point to B} geSCHLagen.

Note that pointing was always inverse to what intonation predicts. In other words, with unstressed pronouns pointing was to (the location in gesture space standing proxy for) the second referent first and to the first referent second, while with stressed pronouns it was to the first referent first and to the second referent second. Hence, pointing was predicted to work in the opposite direction than intonation.

Subjects were shown a video with a person uttering the item in one of the four conditions above. We took care that stress patterns were kept constant over the first two and the last two of the conditions. After having watched the video with the utterance, the subjects were asked to answer a question and choose among two possible answers.

(7) *Wer hat wen geschlagen? (,Who hit whom'?)*

a) *der Peter den Paul (,Paul Peter')*

b) *der Paul den Peter (,Peter Paul')*

Participants were instructed that they have to choose one of the two answers. In case they consider both as possible they should decide for the one that they find more plausible.

We included nine filler items, which involved pointing. They were designed along the lines of the following example.

(8) *Gestern auf der Party hat Peter_{+point to A} Linus_{+point to B} beleidigt. Der_{+point to B} hat dann sofort angefangen zu weinen.*

,Yesterday at the party Peter insulted Linus. He [d-pro] immediately started crying then.'

This is an example where pointing supports the intended meaning. Linus was the one who was insulted and hence the one who most likely started crying. D-pronouns, as in the second sentence, are usually argued to pick up non-subject or non-topical referents (Bosch et al. 2003, Bosch et al. 2007). Hence, in example (8), the d-pronoun is predicted to refer to Linus. Pointing is compatible with this preference, as it is also to (the location associated with) Linus. Second, we also included examples of this kind, but with pointing to the subject so that pointing and d-pronoun resolution are not in line with each other. And third, we had filler items where pointing was misleading with an additional gender mismatch.

(9) *Heute Morgen hat Markus_{+point to A} Lily_{+point to B} einen lang ersehnten Wunsch erfüllt. Die_{+point to A} war danach total aus dem Häuschen.*

,Today Markus fulfilled a long-awaited wish of Lily's. She [d-pro] was totally thrilled afterwards.'

Three items of each sort were included.

Experiment: hypotheses. Our hypothesis was to verify the judgements from the literature and find that people prefer parallel pronoun resolution with unstressed pronouns (condition 1) and inverse pronoun resolution with contrastive stress on the pronouns (condition 2). If, as we argue, co-speech gestures can serve to determine the resolution of pronouns, we expect an effect of gesture in the following way: For condition 3, we expect that people choose inverse pronoun resolution more often than without gesture (in condition 1, cf. results of Goodrich Smith & Hudson Kam 2012 and Foraker 2014). And for condition 4, we predict that, with the contrastive stress pattern, participants favour parallel resolution more often than without gesture (in condition 2). Furthermore, conditions 3 and 4 can tell us more about the pronoun resolution potential of gesture and intonation in comparison to each other. As gesture and intonation work in opposite directions, we can see which of the two has stronger impact. Condition 4 is especially telling in this context. Here, we have strong contrastive stress, which, according to the literature, results in inverse pronoun resolution. On the other hand, gesture marking induces parallel pronoun resolution according to our assumptions. A general preference for inverse resolution would suggest that intonation has a stronger impact than co-speech gesture. Overall preference for parallel resolution would suggest that gesture has a stronger influence than intonation.

Experiment: results³. Data of 16 participants (2 observations per participant per condition) were subjected to a generalized linear mixed model analysis (glmm) with intercepts for participants and items as random factors using the lme4 package in R. The relative frequencies of parallel pronoun resolutions in the four conditions are shown in Figure 1.

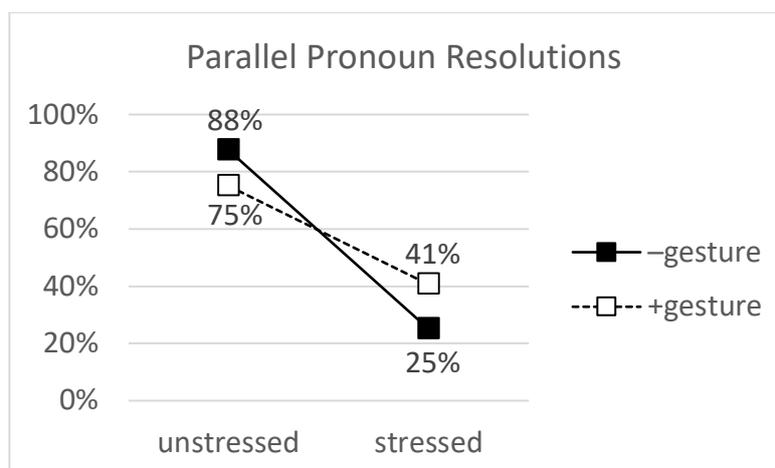


Figure 1. Percentages of parallel pronoun resolutions as a function of intonation and gesture

The analysis confirmed a main effect of intonation ($z = 4.87, p < .001$: 81 % vs. 33 % for stressed and unstressed pronouns). The interaction of intonation with gesture almost reached significance ($z = 1.95, p = .051$). As expected, the gesture in favour of an inverse resolution accompanying the unstressed pronoun decreased the frequency of parallel resolutions, whereas the gesture in favour of a parallel resolution accompanying the stressed pronoun increased the frequency of parallel resolutions. Thus, the statistical interaction corroborates a consistent effect of gesture in the direction that the gesture provides evidence for.

Experiment: discussion. Observations from the literature that unstressed pronouns favour parallel resolution are clearly confirmed (88% answers for parallel resolution). It could also be confirmed that contrastive stress induces inverse pronoun resolution (75% answers for inverse resolution). Furthermore, we found an interaction effect. While there were 88% answers for parallel resolution for unstressed pronouns without gesture, with gesture (working in the opposite direction, i.e. for an inverse resolution) there were only 75%. Also, stressed pronouns

³ I would like to thank Robin Hörnig for help with the statistics and interpretation of the results.

clearly favour inverse readings (only 25% answers for parallel pronoun resolution), but with gestures suggesting the opposite (= parallel) resolution pattern, we find 41% answers with parallel pronoun resolution.

With this experiment we could verify the observations from the literature and we find a clear gesture effect, which means that there is evidence for the claim that gesture can assist to determine pronoun resolution, which confirms results of Goodrich Smith & Hudson Kam 2012 and Foraker 2014. But furthermore, we also observe that intonation has a stronger impact than gesture. Whenever gesture and intonation work in opposite directions, most answers are in line with what would be predicted by the intonation pattern.

Conclusion. After all we can finally ascertain that the Oscar for the most important supporting role in language goes to... – and what else would we dare to present here in this volume? – ... INTONATION (but gesture seems to be more than an extra, too).

References

- Bosch, Peter, Tom Rozario & Yufan Zhao. 2003. Demonstrative pronouns and personal pronouns. German 'der' vs. 'er'. *Proceedings of the EACL 2003*, 61-68.
- Bosch, Peter, Graham Katz & Carla Umbach. 2007. The non-subject bias of German demonstrative pronouns. In: *Anaphors in text*, M. Schwarz-Friesel, M. Consten and M. Knees (eds.), 145-164. Amsterdam & Philadelphia: Benjamins.
- Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. G. Fischer: Jena.
- Büring, Daniel. 1997. The great scope inversion conspiracy. *Linguistics and Philosophy* 20, 175-194.
- Cresti, Diana. 1995. *Indefinite topics*. MIT: Ph.D. dissertation.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Dretske, Fred I. 1972. Contrastive statements. *Philosophical Review* 1972, 411-437.
- Ebert, Cornelia. 2018. A comparison of sign language with speech plus gesture, commentary on Philippe Schlenker's target article "Visible Meaning: Sign Language and the Foundations of Semantics", *Theoretical Linguistics* 44(3-4), 239-250.
- Ebert, Cornelia, Robin Hörnig, Aleksandra Ćwiek, Susanne Fuchs & Manfred Krifka. 2019. Experimental studies on co-speech gestures and their (non-)at-issueness, *poster at LingCologne 2019*, Cologne, June 2019.
<http://www.cow-electric.com/neli/posters/CERHACSFMK-LingCologne2019.pdf>
- Ebert, Cornelia & Ebert, Christian. 2014. Gestures, demonstratives, and the attributive/referential distinction, *slides of talk at Semantics and Philosophy in Europe 7*, ZAS, Berlin, June 2014. <http://www.cow-electric.com/neli/talks/CECE-SPE2014.pdf>
- Endriss, Cornelia. 2009. Quantificational topics – a scopal treatment of exceptional wide scope phenomena. In *Studies in Linguistics & Philosophy*, Springer.
- Endriss, Cornelia & Stefan Hinterwimmer. 2006. Topic interpretation with quantificational adverbs and determiners. In: *Proceedings of NELS 36*, Davis, C., A. Deal, Y. Zabbal (eds.), 241-252. Amherst: GLSA.
- Foraker, Stephani. 2014. Gestures modulate access to referent representations. *CUNY Human Sentence Processing conference*, Columbus, OH.
http://faculty.buffalostate.edu/forakesm/CUNY14-Foraker_ProG_poster.pdf

- Fricke, Ellen. 2009. Deixis, Geste und Raum: Das Bühlersche Zeigfeld als Bühne. In: *Theatralität sprachlichen Handelns. Eine Metaphorik zwischen Linguistik und Kulturwissenschaften*, M. Buss, S. Jautz, F. Liedke und J. Schneider (eds.), 165–186. München: Fink.
- Goodrich Smith, Whitney & Carla L. Hudson Kam. 2012. Knowing ‘who she is’ based on ‘where she is’: the effect of co-speech gesture on pronoun comprehension. *Language and Cognition* 4, 75–98.
- Herburger, Elena. 1993. Focus and the LF of NP quantification. In: *SALT III*, U. Lahiri and A. Z. Wyner (eds.), 77-96. Ithaca: Cornell University, Department of Modern Languages and Linguistics.
- Jäger, Gerhard. 1999. Topic, focus and weak quantifiers. In: *Focus. Linguistic, cognitive and computational perspectives*, P. Bosch and R. van der Sandt (eds.), 187-212. Cambridge: Cambridge University Press.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Krifka, Manfred. 1990. Four thousand ships passed through the lock: object-induced measure functions on events. *Linguistics and Philosophy* 13, 487-520.
- Lakoff, George. 1971. Presupposition and relative well-formedness. In: *Semantics*, D. Steinberg and L. Jakobovits (eds.), 329-340. Cambridge: Cambridge University Press.
- Lascarides, A. & M. Stone (2009): A formal semantic analysis of gesture, *Journal of Semantics* 26(4): 393-449.
- Pafel, Jürgen. 1997. *Skopus und logische Struktur*. Neuphilologische Fakultät, Eberhard-Karls-Universität: Habilitationsschrift.
- Partee, Barbara H. 1991. Topic, focus and quantification. In: *Proceedings from Semantics and Linguistic Theory 1*, Cornell Working Papers in Linguistics, S. Moore and A. Z. Wyner (eds.), 159-188.
- Potts, Chris. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Rooth, Matts E. 1985. *Association with focus*, University of Massachusetts at Amherst: Ph.D. dissertation.
- Schlenker, Philippe. 2018. Gesture projection and cosuppositions, *Linguistics & Philosophy* 41(3), 295–365.
- Syrett, Kristen & Todor Koev. 2015. Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics* 32, 525–577.

Appendix

Items:

1. Erst hat Stefan Charles geärgert. Und dann hat er ihn geschlagen.
2. Erst hat Fritz Pierre eine Uhr geschenkt. Und dann hat er ihn in die Karibik eingeladen.
3. Erst hat Axel Mustafa bei den Hausaufgaben geholfen. Und dann hat er ihm die Abschlussarbeit geschrieben.
4. Erst hat Robert Silas in den Bauch gestupst. Und dann er ihn gebissen.

5. Erst hat Mario Finn das Einmaleins beigebracht. Und dann hat er ihm die Galoistheorie erklärt.
6. Erst hat Moritz Levin ein Buch ausgeliehen. Und dann hat er ihm die komplette Beast-Quest-Kollektion mit nach Hause gegeben.
7. Erst hat Philipp Kevin ausgelacht. Und dann hat er ihn vor der kompletten Schule bloßgestellt.
8. Erst hat Jörg Alessandro in die Pizzeria eingeladen. Und dann hat er ihn in ein Sterne-Restaurant ausgeführt.

Fillers:

1. Gestern auf der Party hat Peter_{+point to A} Linus_{+point to B} beleidigt.
Der_{+point to B} hat dann sofort angefangen zu weinen.
2. Letztes Jahr hat Sven_{+point to A} Hans_{+point to B} beim Tennis geschlagen.
Der_{+point to A} forderte sofort eine Revanche.
3. Heute Morgen hat Markus_{+point to A} Lily_{+point to B} einen lang ersehnten Wunsch erfüllt.
Die_{+point to A} war danach total aus dem Häuschen.
4. Neulich hat Henry_{+point to A} Tom_{+point to B} das Handy geklaut.
Der_{+point to B} suchte erfolglos den ganzen Tag danach.
5. Gestern hat Rob_{+point to A} Sam_{+point to B} mit einem Ausflug ins Phantasialand überrascht.
Der_{+point to A} wollte da schon immer mal hingehen.
6. Letztes Wochenende hat Daniel_{+point to A} Laura_{+point to B} im Schwimmbad ins Wasser geschubst.
Die_{+point to A} hatte das überhaupt nicht kommen sehen.
7. Vorhin hat Karl_{+point to A} Leonard_{+point to B} einen Stift geliehen.
Der_{+point to B} vergisst nämlich ständig sein Federmäppchen.
8. Vor ein paar Tagen hat Michael_{+point to A} Lukas_{+point to B} beim Umzug geholfen.
Der_{+point to A} war ihm dafür sehr dankbar.
9. Letzte Woche hat Thomas_{+point to A} Lea_{+point to B} beim Fußballspielen verletzt.
Die_{+point to A} musste daraufhin ausgewechselt werden.