Alignment in Iconic Gestures

Mol, L.; Krahmer, E.J.; Swerts, M.G.J.

Published in:
Proceedings of the eight International Conference on Auditory-Visual Speech Processing (AVSP 2009)

Publication date:
2009

Alignment in Iconic Gestures: Does it make sense?

Lisette Mol, Emiel Krahmer and Marc Swerts

Communication and Cognition group, Department of Humanities, Tilburg University, The Netherlands
{L.Mol, E.J.Krahmer, M.G.J.Swerts}@uvt.nl

Abstract

Many studies have shown that people imitate and repeat each other's behaviors. This holds for both verbal and nonverbal behavior. The production of co-speech hand gestures is a special case of nonverbal behavior, because it is believed to be tightly linked to verbal language production and because gestures can carry meaning in a way that is similar to verbal language. It has been shown that people reuse each other's hand shapes for co-speech gestures. This study looks at the relevance of gestures' meaning and their relation to speech for such mimicry to occur. In two studies we found that speakers mimicked iconic gestures that they had observed, but only if these gestures were consistent with the co-occurring speech. This is evidence that the mimicry of iconic gestures is an instance of convergence of linguistic behavior, also known as *alignment*, rather than a supercial imitation of physical behavior.

Index Terms: alignment, gesture, narration

1 Introduction

People frequently mimic each other's behavior. Mimicry is defined as one person repeating the behavior of another person [1]. Some forms of mimicry can have a clear purpose, as they enable one to learn important functional behaviors [2]. It has also been found that mimicking others can have significant social benefits. Van Baaren et al. [3] found that a waitress received higher than usual tips when repeating her customers' orders literally. Yet for some instances of mimicy it seems less obvious what the social or functional validity could be [4]. For example, if one person starts yawning, oftentimes those around will start yawning as well. The same holds for rubbing one's chin. Chartrand and Bargh explain this type of behavior with the *perception-behavior link*, meaning that "the mere perception of another's behavior automatically increases the likelihood of engaging in that behavior oneself", [4] p. 893. Importantly, they state that although such mimicry may act as a kind of 'social glue', intent or conscious effort are not required for this phenomenon to occur.

When in dialogue, people's behavior converges at many different levels [5]. Branigan et al. use the term *alignment* to refer to the convergence of linguistic behavior. For example, if two people each refer to a certain type of seat with the word 'sofa', they are aligned on the use of this term. Whereas if one person uses 'couch' and the other uses 'sofa' they are not. Pickering and Garrod [6] originally used *alignment* to refer to a similarity in mental representations rather than behavior. Thus, if one person would use the word 'sofa', leading another person to merely think of this type of seat as a 'sofa', they would also be aligned.

Pickering and Garrod stress that while in dialogue, interlocutors align their representations at many different levels, such as the lexical and syntactic level. They also state that alignment at one level can enhance alignment at other levels, since the mental representations at many different levels are linked bidirectionally. They propose alignment is a process which is direct and automatic, using priming as a mechanism, much like social behavior following the perception-behavior link. However, in linguistic alignment sometimes the meaning of the mimicked behavior is highly relevant. For example, Costa et al. [7] found that although people frequently align on the term 'wheel' to refer to a tire when interacting with a second language learner, they are far less likely to do so when interacting with a native speaker.

As examples of linguistic alignment, people have been shown to reuse each other's (referring) expressions, also known as *lexical entrainment*, (i.e. [8], [9]) and also each other's syntactic constructions (i.e. [10], [11]). These phenomena have been shown to occur in a non-dialogue context as well, such as when observing others [12], or repeating one's own constructions or words when doing a task individually [1], [13]. This indicates that speaker-centered processes may be involved. Additionally, some studies propose that apart from direct and automatic processes being involved in alignment, audience design may play a role as well [5], [13]. For example, Branigan et al. found that people were more likely to repeat words that a computer proposed, rather than maintaining their own referring expression, if they had less confidence in the computer's performance. Thus, like for other forms of behavioral mimicry, there may be instances of linguistic alignment for which the (social) function is either more or less clear.

Besides verbal mimicry, people have also been shown to mimic many other behaviors [14], [15], [4], such as self-adaptors, laughter [16], facial expressions [17], [18], emotions [19], and mood [20]. Our present study focusses on the mimicry of hand gestures that people spontaneously produce while speaking. These gestures are not as symbolic or conventional as verbal language is, but rather are more idiosyncratic [21]. Yet co-speech hand gestures are thought to have a close relationship with speech [22], [21]. For example, certain simple biphasic hand movements can be used to emphasize certain parts of speech [23]. Other gestures, called *illustrators* or *iconic gestures* are thought to express part of the content that a speaker is trying to convey verbally as well. For example, when someone talks about a person throwing away a heavy object, an iconic gesture could display how the object was held, how big it was, with what velocity it was thrown away and in what direction, etc. It has been shown that such gestures, like speech, can convey meaning to an addressee as well, i.e. [24].

1.1 Mimicry of co-speech hand gestures

For certain co-speech hand gestures, instances where the same encoding patterns were repeated across speakers have been

found [25], [26], [27]. Kimbara [28] showed that in a joint description task, the hand shapes with which iconic gestures were produced were more similar when interlocutors could see each other than when they were separated by an opaque screen. In both of these settings participants could hear each other, so they could align verbally. But only in the condition without a screen could participants see each other's gestures, enabling them to mimic each other's gestures as well, and they did. This indicates that one speaker's gestures influence another speaker's gestures directly, rather than interlocutors' gestures resembling each other as a result of verbal alignment. But the question remains open whether gestural mimicry is a case of automatically engaging in a behavior that one perceives, regardless of its meaning or functional validity, or whether it is tied to alignment in representations, which may be similar or linked to the representations underlying linguistic alignment.

In [12] it was shown that gestural mimicry can also be found outside of dialogue. Participants watched a movie clip in which interlocutors were conferring on a route to take through a certain model town in front of them. The interlocutors were seen from the back, which allowed for manipulation of this stimulus movie. Two different video images, in which interlocutors either mimicked each other's gestures or not, were combined with two different audio tracks, in which interlocutors either mimicked each other's speech or not, rendering a total of four stimulus movies. It was found that participants who had seen or heard more mimicry, produced more of the mimicked features while subsequently describing the stimulus movie to the experimenter. This held both for gestural as well as verbal mimicry. However, interestingly, no significant interactions were found between verbal and gestural mimicry. So participants that had seen the movie clip in which interlocutors mimicked each other's gesture and speech did not mimic the interlocutors' gestures more than did participants that had seen the clip in which the interlocutors solely mimicked each other's gestures.

This goes somewhat against a prediction made by Pickering and Garrod [6] that alignment at one level can facilitate alignment at another level, because of bidirectional connections between each level's representations. Especially since gesture and speech are often believed to be linked to the same, or linked, underlying representations i.e. [22]. For example, Krauss [29] found evidence that gesturing can sometimes aid speakers to retrieve lexical forms. And De Ruiter [30] proposed that gestures can help a speaker in maintaining a mental image while it is being verbally expressed. So if people would align their mental representations, it could be expected that lexical alignment and gestural mimicry would correlate. But this seems contradictory to the evidence found so far. Thus, again the question remains open whether mimicry in gesture is tied to underlying cognitive representations.

1.2 Present study

In this study we want to address the question whether mimicry of hand gestures is a case of automatically engaging in a behavior one perceives, even if its meaning is not apparent, or a case of mimicry that may be based on alignment of representations, like linguistic alignment can be. As explained above, gestures are sometimes thought to be linked to the mental representations of a speaker. If gestures contribute to the mental representations of comprehenders as well, it is to be expected that gestures that match a speaker's verbal description differ fundamentally from gestures that are inconsistent with the co-occurring speech. The mimicry of wildly inconsistent gestures would most likely be an example of the perception-behavior link, since their meaning is not very relevant to the task at hand. But the mimicry of consistent gestures on the other hand may be a case of linguistic alignment. Their meaning may activate representations that are consistent with the representations activated by speech, or may even be integrated into those representations.

Our first study tests our method for eliciting mimicry of iconic gestures. We investigate whether participants who see certain iconic gestures in a stimulus movie are more likely to subsequently display these gestures when retelling the stimulus to an experimenter. Our second study compares gestures that are consistent with the co-occurring speech to gestures that are inconsistent with the co-occurring speech. In addition to whether or not each of these gesture types will be mimicked, we are also interested in the timing aspect of gestural mimicry. If speech and gesture give rise to a single mental concept, the mimicked gesture is expected to co-occur with a verbal description of the same concept as in the stimulus movie. Thus, if the meaning of the gesture in relation to the meaning of the concurrent speech is relevant to gestural mimicry, the timing of the mimicked gesture may be more critical than when it is a case of merely mimicking a physical behavior. Additionally, if seeing a gesture contributes to the mental concept that is subsequently being expressed, it may be more likely that this concept is again expressed with a combination of speech and gesture.

2 Study I

2.1 Participants

38 participants (28 female) volunteered for this study. They were all native speakers of Dutch and most of them were students of Tilburg University.

2.2 Stimuli

For this study two different movie clips were created. In both movie clips the same male speaker told the story of an animated cartoon ('Canary Row' by Warner Brothers) as though he had just watched it. The movie clip consisted of ten fragments. It started with a short introduction in which the speaker stated that the cartoon was a Tweety and Sylvester movie in which Sylvester (a cat) tries to capture Tweety (a pet bird). Then followed eight fragments each describing one episode of the cartoon, which corresponds to one attempt of Sylvester to catch Tweety. These fragments were about 15 seconds long. The last fragment consisted of a short closure. Blank video was inserted in between the fragments, to allow the movie to be paused at the appropriate times. The speaker was seated on a chair and looked straight into the camera. The image showed the entire upper-body of the speaker.

The two versions differed in the number of iconic hand gestures that the speaker produced. In one version, an iconic gesture describing an action was produced for each episode of the cartoon. These gestures were based on retellings of participants in a previous study [31]. They consisted of:

binoculars two hands (cylinder shaped) are held in front of the eyes, as though looking through binoculars

drain pipe two hands/ arms make climbing/ grabbing motions

while moving upwards

rolling ball two hands spin around each other from the wrists

money tin right hand imitates the holding and shaking of a money tin

creeping hands (flat, palms down) and arms are moved forward one by one, imitating a creeping motion

throwing the weight two hands (fingers spread, palms facing each other) are held about 30 cm apart, while a motion is made starting at the head and moving forward in an arc, as though throwing something big away from oneself

swinging two hands are held on top of each other and quickly make a grasping motion above the head

running arms are moved as while running, close to the body of the speaker

In the other version no iconic gestures were produced. No other hand gestures were produced in any of the two versions and body posture, voice quality, facial expressions, and other prosodic factors were kept constant across versions. In both versions, the speaker used eight target phrases. These were unusual ways of putting things, for example 'as a full-blown Tarzan' (Dutch: als een volleerde Tarzan) or 'the yearly spring call of the canary' (Dutch: de jaarlijkse lenteroep van de kanarie). These target phrases were the same in both versions and they never occurred simultaneously with a target gesture.

Participants only saw one of these two stimulus movies of a speaker retelling the original cartoon movie and did not see the animated cartoon themselves.

2.3 Design

We have used a between subjects design with two conditions. The 'Gestures condition' differed from the 'No gestures condition' in that the speaker in the stimulus movie produced one iconic gesture per fragment, rather than no iconic gestures at all. In each condition the stimulus movie contained certain target phrases for measuring lexical entrainment.

2.4 Procedure

Participants came to the lab and were randomly assigned to the Gestures or No gestures condition. They read the instructions, which explained the task as a memory task, in which they had to watch video fragments of a speaker telling a story and subsequently retell these story fragments to the experimenter. Participants were instructed to take their time when retelling the stories. They were given the opportunity to ask further clarification and once all was clear the experiment started.

Participants first watched the introductory fragment, which they did not have to retell. Then they watched the fragments describing the cartoon episodes one at a time. After each fragment, participants paused the movie and turned their office chairs ninety degrees such that they were facing the experimenter while they retold the story. A camera was placed to the side of the experimenter, recording the participants' narrations (participants were told they were videotaped to facilitate our analyses afterwards). The experimenter did not interrupt the participants and did not produce any hand gestures, but did show other non verbal signs of listening to their story in a natural way. Finally participants watched the last fragment, which they did not have to retell. All participants gave written consent for their data to be used for scientific purposes.

2.5 Coding & Analysis

We coded all gestures that matched gestures in the stimulus movie of the Gestures condition as *Target gesture*. Gestures that did not match completely but matched the target gesture in either the place and shape of the hand(s) or the direction in which the gesture was made were labeled as *Partial Target Gesture*. If no gesture or a different gesture was produced at the place in the narration of the original target gesture this was labeled as *No or Different gesture*.

If the full target phrase was used this was labeled as *Full* verbal alignment. If one or more words of the target phrase were reproduced this was labeled as *Partial* verbal alignment. One exception has been made in this respect. The target phrase of the third episode consisted of five content units (shiny blue Chiquita banana emblem, Dutch: blinkend blauw Chiquita banaan embleem) which was exceptionally many. In this case reproduction of four or more of these words has been counted as full alignment.

Participants that deviated more than three standard deviations from the mean in the total number of iconic gestures produced were excluded from our analysis. This left 18 out of 19 participants in both conditions.

2.6 Results

A Pearson chi-square test showed that the different categories of gestures occurred with different frequencies in each condition. Target gestures were more frequent in the condition in which participants had seen the speaker produce these target gestures (17 vs. 2), see Table 1. Comparison of means showed that the effect of condition on the average total number of iconic gestures that participants produced was not significant, $p = .14$. Unlike mimicry of gestures, verbal alignment occurred equally often in the Gestures and No Gestures condition, see Table 2. Figure 1 shows an example of a target gesture being repeated by a participant.

Table 1: Overview of gestures at target moments

Condition:	Target Gesture	Partial Target Gesture	No / Different Gesture	Total
No gestures	2 (1.4%)	9 (6.3%)	133 (92.4%)	144 (100%)
Gestures	17 (11.8%)	9 (6.3%)	118 (81.9%)	144 (100%)
Total	19 (6.6%)	18 (6.3%)	251 (87.2%)	288 (100%)

$\chi^2(2) = 12.74,\ p < .01$

2.7 Discussion

This study was able to elicit mimicry of iconic gestures. Certain iconic gestures occurred more frequently if participants had seen these gestures in the stimulus movie. Since the difference between the two conditions in the average total number of iconic gestures produced was not significant, it seems that this was not simply an effect of participants in the Gestures condition gesturing more

Table 2: Verbal alignment

Condition:	Full	Partial	No	Total
No gestures	27 (18.8%)	28 (19.4%)	89 (61.8%)	144 (100%)
Gestures	27 (18.8%)	29 (20.1%)	88 (61.1%)	144 (100%)
Total	54 (18.8%)	57 (19.8%)	177 (61.5%)	288 (100%)

$\chi^2(2) = .02\ n.s.$

Stimulus movie Participant

Figure 1: Example of mimicry of an iconic hand gesture (binoculars).

and therefore being more likely to produce the target gestures. Rather, the form of the specific gestures was taken over. But is this copying of the form of a gesture merely an instance of imitating the speaker's body movement, or is the meaning of the gesture relevant to this process? Our second study addresses this question by varying the relation between the speaker's gestures and speech.

3 Study II

3.1 Participants

47 participants (33 female) volunteered for this study. They were all native speakers of Dutch and most of them were students of Tilburg University. None of the participants had participated in study I.

3.2 Stimuli

For this study again two movie clips were produced, which were very similar to the clips in study I. The first movie was made in the same way as the clip containing iconic gestures in study I. In the second movie clip, the speaker produced one iconic gesture per episode as well, however this time the gesture did not match the speaker's verbal description. An iconic gesture from another episode was produced instead of the original gesture, at the same place in the verbal description. For example, instead of the binoculars gesture, the speaker produced the running gesture while actually verbally referring to an event which included binoculars. Still, at first sight this movie clip looked as natural as the other one. The same speaker was used as in study I and all other factors, including the target phrases, were kept constant.

3.3 Design & Procedure

We have used a between subjects design with two conditions. The 'Congruent' condition differed from the 'Incongruent' condition in that the speaker in the stimulus movie produced iconic gestures that matched his co-occurring verbal description, rather than iconic gestures that did not match his co-occurring verbal description. The procedure was the same as for study I. When asked by the experimenter, none of the participants showed any indication of suspecting that the experiment was about the mimicry of gestures.

3.4 Coding

Each target gesture in the stimulus movie occurred at a target moment: a given place in the verbal narration. We looked at participants' iconic gestures at those target moments. We coded gestures that matched the corresponding gesture in the stimulus movie of the participant's condition as *Target Gesture Own Condition* and gestures that matched the corresponding gesture from the stimulus movie of the other condition as *Target Gesture Other Condition*. Gestures that matched the corresponding target gesture of the Congruent condition in either the place and shape of the hand(s) or the direction in which the gesture was made were labeled as *Partial Target Gesture Congruent* (There were no partial matches with corresponding gestures from the stimulus movie of the Incongruent condition.) If a different gesture was produced at the moment of the original target gesture this was labeled as a *Different gesture*. Verbal alignment was coded in the same way as in study 1.

3.5 Results & Discussion

A Pearson chi-square test showed that the different categories of gestures occurred with different frequencies in each condition, see Table 3. In the Congruent condition, far more target gestures of the participant's own condition were produced than in the incongruent condition (19 vs. 1)[1]. There were four instances in which a target gesture from the Congruent condition was produced in the Incongruent condition. Most likely this is because participants produced these gestures spontaneously.

In the Congruent condition, it was more likely that a gesture was produced at a target moment than in the Incongruent condition (71 vs. 39), Yates $\chi^2(1) = 10.56\ p < .001$. No incongruent gestures were partially reproduced. We found no significant difference between the two conditions in the average total number of iconic gestures that participants produced, $p = .16$.

With a Pearson chi-square test, we found that our categories for verbal alignment showed no significant difference in frequencies between the two conditions, see Table 4. Comparisons of means showed that full target phrases were reproduced equally often by participants in the Congruent and Incongruent condition ($p = .79$), but participants in the Congruent condition ($M = 1.92, SD = 1.1$) produced more partial target phrases than participants in the Incongruent condition ($M = 1.26$, $SD = .92$), $t(45) = 2.27, p < .05$.

[1] In the Congruent condition, 11 participants produced at least one target gesture and the maximum number of target gestures produced by a participant was 4. Comparison of means, $t(45) = 2.540, p < .02$, also showed that participants in the Congruent condition ($M = .79, SD = 1.1$) produced more target gestures than participants in the Incongruent condition ($M = .17, SD = .39$).

Table 3: Overview of gestures at target moments

Condition:	Target Gesture: Own Condition	Target Gesture: Other Condition	Partial Target Gesture Congruent	Different Gesture	No Gesture	Total
Incongruent	1 (.5%)	4 (2.2%)	13 (7.1%)	21 (11.4%)	145 (78.9%)	184 (100%)
Congruent	19 (9.9%)	0 (0%)	23 (12.0%)	29 (15.1%)	121 (63.0%)	192 (100%)
Total	20 (5.3%)	4 (1.1%)	36 (9.6%)	50 (13.3%)	266 (70.7%)	376 (100%)

$\chi^2(4) = 26.26,\ p < .0001$[2]

Table 4: Verbal alignment

Condition:	Full	Partial	No	Total
Incongruent	25 (13.6%)	29 (15.8%)	130 (70.7%)	184 (100%)
Congruent	27 (14.1%)	46 (24.0%)	119 (62.0%)	192 (100%)
Total	52 (13.8%)	75 (19.9%)	249 (66.2%)	376 (100%)

$\chi^2(2) = 4.15,\ n.s.$

This study has shown that the reproduction of iconic gestures found in study I is not merely an instance of copying the speaker's physical behavior. Gestures that did not match the speaker's speech were almost not repeated by participants. Participants who had seen 'nonsense' gestures even made fewer gestures at the target moments than participants who had seen meaningful gestures. Therefore, the copying of a gesture's form is more likely a case of convergence in linguistic behavior (alignment) than a general instance of physical mimicry.

4 General Discussion & Conclusion

Participants reproduced iconic gestures that they had seen. Our first study shows that participants who had seen a speaker perform certain iconic gestures while retelling an animated cartoon were more likely to use the same gestures when asked to retell this story to a third person. Our second study shows that this process is not just imitation of the speaker's physical behavior. Only gestures that were consistent with the speaker's verbal description, and thus carried relevant meaning, were reproduced. The reuse of the form of an iconic gesture of another speaker therefore seems to be a case of convergence of linguistic behavior: alignment.

Participants that had seen a matching gesture were also more likely to produce a gesture at the point where the original target gesture had been, than were participants who had seen a gesture that did not match the speech in any way. This may mean that rather than making a gesture that completely did not align with the original speaker, participants that saw non matching gestures preferred not making a gesture at all at that place. Or that participants in the Congruent condition not only aligned their gestures in their form, but also in their placing. This would be consistent with a social explanation of mimicry (i.e. [4], [3]).

However, it could also be that the mismatching gesture had distorted participants' mental imagery, such that there wasn't a clear enough mental image to base their gesture on. Or that the combination of speech and a matching gesture had given rise to a representation that was especially suitable for expression in speech and gesture (but not necessarily the exact same speech and gesture that it originally occurred with). Additional studies are needed to investigate the relation between alignment in gesture and alignment in mental representations. Our study indicates that they may be linked.

In the Congruent condition, in which more target gestures were (partially) reproduced, more target phrases were partially reproduced as well. But remember that these phrases and gestures did not occur simultaneously. Therefore this is more likely explained by participants aligning less with a less coherent speaker, than by alignment in expressions and gesture enhancing each other. The absence of a difference in the extent to which verbal alignment occurred in the Gestures and No Gestures condition in study I also points in this direction.

Clearly, this study was able to elicit some alignment. Both target phrases and target gestures were reproduced reliably, provided that they made sense. However, this study also has its limitations. There was no interaction between the speaker and the participant and the story was retold to a relative outsider. In a future study we want to study gesture in a setting that is closer to natural face-to-face interaction. We would also like to examine the effects of more subtle mismatches between speech and gesture and to further investigate whether alignment in gestures is driven by alignment in mental representations. This we are planning to do with an interactive route description task.

5 Acknowledgements

We thank Els Jansen and Anouk van den Berge for collecting and coding the data of the first and second experiment respectively. We also like to thank all students that took part in these studies.

References

[1] J. Bock, "Syntactic persistence in language production," *Cognitive Psychology*, vol. 18, pp. 335–387, 1986.

[2] M. Tomasello, E. Savage-Rumbaugh, and A. Kruger, "Imitatative learning of actions on objects by children, chim-

[2] In this test, 2 of the expected cell frequencies are smaller than 5. Therefore, we repeated the analysis with gestures that matched target gestures of the other condition counted as 'Different': $\chi^2(3) = 21.28,\ p < .0001$.

panzees, and enculturated chimpanzees," *Child Development*, vol. 64, pp. 1688–1705, 1993.

[3] R. Van Baaren, R. Holland, B. Steenaert, and A. Van Knippenberg, "Mimicry for money: Behavioral consequences of imitation," *Journal of Experimental Social Psychology*, vol. 39(4), pp. 393–398, 2003.

[4] T. Chartrand and J. Bargh, "The chameleon effect: The perception-behavior link and social interaction," *Journal of Personality and Social Psychology*, vol. 76, pp. 893–910, 1999.

[5] H. Branigan, M. Pickering, J. Pearson, and J. McLean, "Linguistic alignment between people and computers," *Journal of Pragmatics*, In Press.

[6] M. Pickering and S. Garrod, "Toward a mechanistic psychology of dialogue," *Behavioral and Brain Sciences*, vol. 27, pp. 169–226, 2004.

[7] Costa, Albert, Pickering, Martin, Sorace, and Antonella, "Alignment in second language dialogue," *Language and Cognitive Processes*, vol. 23(4), pp. 528–556, June 2008.

[8] H. H. Clark and D. Wilkes-Gibbs, "Referring as a collaborative process," *Cognition*, vol. 22, pp. 1–39, 1986.

[9] S. Garrod and A. Anderson, "Saying what you mean in dialogue: A study in conceptual and semantic co-ordination," *Cognition*, vol. 27, pp. 181–218, 1987.

[10] W. Levelt and S. Kelter, "Surface form and memory in question answering," *Cognitive Psychology*, vol. 14(1), pp. 78–106, 1982.

[11] H. Branigan, M. Pickering, and A. Cleland, "Syntactic coordination in dialogue," *Cognition*, vol. 75 (2), pp. B13–B25, 2000.

[12] F. Parrill and I. Kimbara, "Seeing and hearing double: The influence of mimicry in speech and gesture on observers," *Journal of Nonverbal Behavior*, vol. 30, pp. 157–166, 2006.

[13] M. Van Der Wege, "Lexical entrainment and lexical differentiation in reference phrase choice," *Journal of Memory and Language*, vol. 60, pp. 448–463, 2009.

[14] T. Chartrand, W. Maddux, and J. Lakin, *Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry.* New York: Oxford University Press, 2005, pp. 334–361.

[15] M. LaFrance, "Nonverbal synchrony and rapport: Analysis by the cross-lag panel technique," *Social Psychology Quarterly*, vol. 42(1), pp. 66–70, 1979.

[16] R. Young and M. Frye, "Some are laughing; some are notwhy?" *Psychological Reports*, vol. 18, pp. 747–752, 1966.

[17] J. Bavelas, C. Lemery, and J. Mullet, ""i show how you feel." motor mimicry as a communicative act," *Cognitive Psychology*, vol. 50, pp. 322–329, 1986.

[18] E. Hsee, J. Hatfield, J. Carlson, and C. Chemtob, "The effect of power on susceptibility to emotional contagion," *Cognition and Emotion*, vol. 4, pp. 327–340, 1990.

[19] E. Hatfield, J. Cacioppo, and R. Rapson, *Emotional contagion*. Cambridge University Press, 1994.

[20] R. Neumann and F. Strack, ""mood contagion": The automatic transfer of mood between persons," *Journal of Personality and Social Psychology*, vol. 79, pp. 211–223, 2000.

[21] A. Kendon, *Visible action as utterance*. Cambridge University Press, 2004.

[22] D. McNeill, *Hand and Mind: what gestures reveal about thought*. University of Chicago Press, 1992.

[23] E. Krahmer and M. Swerts, "The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception," *Journal of Memory and Language*, vol. 57(3), pp. 396–414, 2007.

[24] G. Beattie and H. Shovelton, "Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech," *Journal of Language and Social Psychology*, vol. 18, pp. 438–462, 1999.

[25] I. Kimbara, "On gestural mimicry," *Gesture*, vol. 6(1), pp. 39–61, 2006.

[26] M. de Fornel, *The return gesture*. Amsterdam: John Benjamins, 1992, pp. 159–193.

[27] A. Tabensky, "Gesture and speech rephrasings in conversation," *Gesture*, vol. 1(2), pp. 213–235, 2001.

[28] I. Kimbara, "Gesture form convergence in joint description," *Journal of Nonverbal Behavior*, vol. 32, pp. 123–131, 2008.

[29] R. Krauss, "Why do we gesture when we speak?" *Current Directions in Psychological Science*, vol. 7, pp. 54–60, 1998.

[30] J.-P. De Ruiter, "Gesture and speech production," Ph.D. dissertation, University of Nijmegen, 1998.

[31] L. Mol, E. Krahmer, F. Maes, and M. Swerts, "The communicative import of gestures: evidence from a comparison of human-human and human-machine interactions," *Gesture*, vol. 9(1), pp. 97–126, 2009.