Temporal Aspects of Behavioral Alignment in Collaborative Remembering

Lucas Bietti (lucas@bietti.org)
Institute for Advanced Study in the Humanities (KWI), Essen, Germany

Kasper Kok (k.i.kok@vu.nl)
Department of Language and Communication, VU University, Amsterdam, Netherlands

Alan Cienki (a.cienki@vu.nl)
Department of Language and Communication, VU University, Amsterdam, Netherlands

Abstract

The coordination of verbal and non-verbal facets of communication between interlocutors appears to be one of the basic cognitive tuning processes for social interaction. In this paper we examine the temporal aspects of behavioral alignment in small group interactions that take place in a natural setting. We find that participants tend to align their body posture and gestures in a sequential rather than simultaneous manner. Our results furthermore suggest that behavioral resonance generally happens fast, but can also occur with substantial delay.

Keywords: everyday activities, gestural alignment, small groups, collaborative remembering

Introduction

The extent of many everyday activities, from brief conversational exchanges about the weather to cooperative assembly of IKEA furniture, reaches beyond the sum of the individuals involved. Rather, many such activities can be seen as *joint actions*: "a social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment" (Knoblich, Butterfill & Sebanz, 2011: 60).

In dialogue, a particularly complex type of joint action, coordination of behaviors amongst participants has been examined on various levels of linguistic organization, from words to choice of sentence structure (e.g., Allen et al. 2011; Fusaroli et al., 2012; Pickering & Garrod, 2004). Adaptation to the linguistic behaviors of others, according to some, contributes to the cognitive accessibility of linguistic resources (Pickering & Garrod, 2004), or to the construction and validation of a *common ground* (Clark & Brennan, 1991).

In close connection with this research, Du Bois (2001) proposed a *dialogic syntax*, focusing on "structural similarities between immediately co-present segments in a broader conversational context" (Du Bois, 2001: 2). Essential to Du Bois' work is the observation that structural similarities between dialogical utterances occur beyond the sentence level and are not necessarily bound to immediately consecutive dialogical turns. The precise temporal characteristics of such linguistic parallelism in conversation, however, have as of yet not been comprehensively quantified.

Somewhat independently of these works, a number of studies in the past decades have examined patterns of coordination between speakers in non-verbal domains. Empirical studies have suggested that when engaged in conversation, people tend to align features of prosody (Levitan & Hirschberg, 2011), eye gaze (Richardson, Dale and Kirkham, 2007), body posture (De Fornel, 1992; Shockley et al., 2007) and gesture (Furuyama, 2000).

Only recently have the temporal dynamics of behavioral resonance received serious attention (Richardson, Dale and Shockley, 2008; Louwerse et al., 2012). Louwerse and colleagues studied the interplay of a range of verbal and non-verbal behaviors in an experimental setting where participants were involved in a route-communication task. Their results suggest that in task-driven interaction, people tend to match different types of behaviors in a synchronized fashion.

In the current paper, we examine the time course of behavioral alignment in group conversations that take place in a more natural environment.

Methods

Participants

Two groups of four participants were recruited for this study via one of the researcher's social networks in Argentina. One consisted of close friends (aged 21 to 23), the other of family members (parents aged 57 and 63, and children aged 30 and 32). Both groups had gone through a shared event together (e.g. a vacation) and had carried along a set of pictures related to that event. All were right-handed and native Spanish speakers.

Procedure and data collection

The data were collected between August and October 2011 in Buenos Aires. Two group conversations were recorded in the homes of one of the participants of each group.

Participants were asked to recollect memories based on the pictures they had brought along. The conversation among family members was about a trip to Maui, Hawaii, that they had made together to attend the wedding of the

parents' daughter. The conversation among the friends was about a two-week summer vacation that they had spent together a in summer resort by the seaside. The interventions of the research assistant were minimal, merely functioning to invoke continuation of the conversation.

The conversations were recorded with two digital video cameras placed in different corners of the room.

Coding

The audio and video recordings were transcribed in detail using two specialized software programs (Inqscribe and ChronoviZ). The overall length of the recordings was 14.54 minutes for the family, and 22.08 minutes for the friends. Video and audio recordings of both focus groups were coded for a wide range of co-verbal behaviors, divided into three categories: (i) manual gestures, further divided into deictic gestures (pointing) and non-deictic gestures (all other types); (ii) non-manual gestures, including shoulder shrugs and head nods; (iii) postural behaviors, including direction of leaning and position of the hands relative to the body. Notably, these categories emerged from the data and were

not imposed by the researchers' predictions made beforehand.

The minimum unit of time used for coding of the video and audio recording was 500ms. For all time points, a binary value was assigned to each behavior for each participant, indicating whether he or she performed that behavior at that particular moment in time.

Analysis

A qualitative example

The following example illustrates some ways in which alignment on lexical, syntactic and gestural levels provides structure to the conversational discourse during collaborative remembering. In the discourse fragment displayed in figure 1, the family members recollect their memories regarding a specific activity – a trip to the beach they had made during the first day of their vacation. The collaborative negotiation of this memory goes along with repeated use of particular syntactic patterns (e.g. 'fuimos a la playa'/'we went to the beach') and lexical items (e.g. 'playa'/ 'beach', 'sí'/ 'yes').
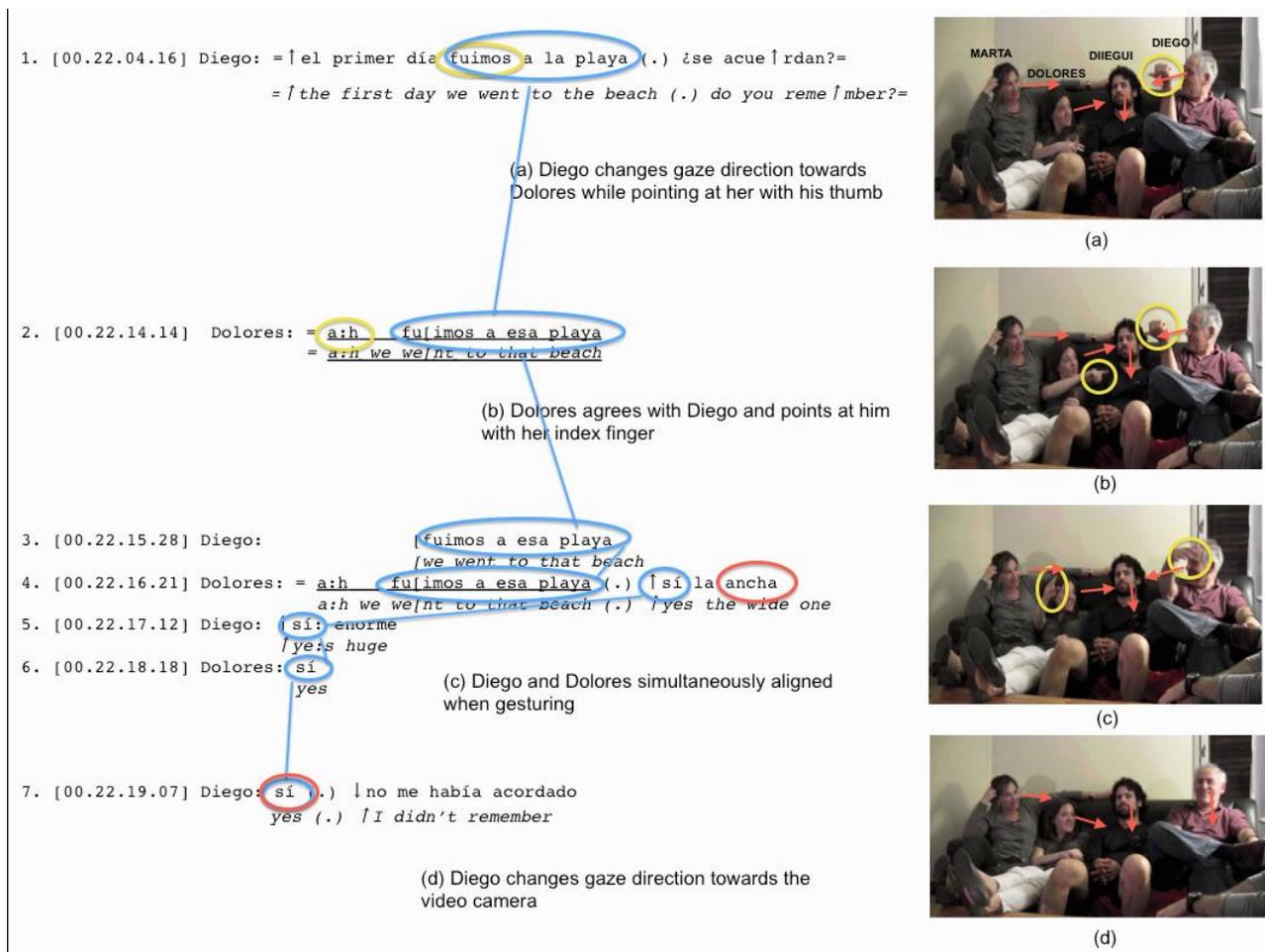

Figure 1: Fragment of the recordings of the family participant group

During this same fragment, two of the participants (Diego and Dolores) first point at each other (b) and then concurrently perform a representational gesture to indicate how wide the beach was (c). In keeping with the research discussed above, these observations seem to show a case where lexical, syntactic and gestural coordination helps to establish common ground and aid memory retrieval for the individual (Diego's memory of the event he had forgotten). Collaborative, multimodal reconstruction of joint activities thus provides a rich resource for accessing informational details which can augment an individual's memory.

Though illustrative, this example is not itself generalizable. Moreover, it does not preclude the possibility that the assumed causal relation between the two gestures is merely an outcome of chance. To address these pitfalls, we have carried out a number of additional analyses of the data from a quantitative perspective.

Quantitative analyses

First, we addressed the question of whether simultaneous and sequential alignment of the behaviors coded occurred more often than predictable on the basis of chance. Note that when members of group of people act randomly (i.e. without any regard to one another) for a long enough time, it is probable that some ostensible instances of 'behavior matching' occur by mere coincidence.

In order to compare our observations with what could be expected by chance, we computed a random baseline via stratified resampling procedure: for all participants, the order of observed sequences of behaviors was shuffled one thousand times, keeping the sequences themselves intact, to yield a distribution of virtual data. Our conclusions are based on whether or not observed alignment rates fall within the 95% most probable alignment rates yielded by the chance simulations. We make no assumptions about the underlying distribution on the population level and therefore do not report p-values.

Two types of alignment were distinguished. Time points where a given behavior (e.g., pointing) was simultaneously performed by two or more participants were qualified as instances of 'simultaneous alignment'. Time points where a

behavior was initiated within ten seconds after another participant had withdrawn that same behavior were counted as instances of 'sequential alignment' (this limit was motivated by the outcome of the second analysis, below).

For all coded behaviors, we compared instances of simultaneous and sequential alignment, defined as such, to the chance baseline. For the purpose of the present paper, we accumulated the alignment counts of the individual behaviors into three clusters (see the coding section).

As figure 2 shows, sequential alignment occurs systematically throughout the course of the conversation for all three of the behavior categories inspected – although the significance is marginal for the non-manual gestures. Interestingly, significant simultaneous alignment is not observed. For manual and non-manual gestures, there is in fact a trend in the opposite direction: simultaneous behavioral coordination happens less often than expectable by chance. The most plausible explanation for this finding is that gestural behaviors are closely linked to speech production processes, and therefore relatively unlikely to be observed in multiple participants at the same time. An alternative, more speculative explanation for this effect is that 'disalignment' serves a social function, e.g. for purposes of signaling disengagement or maintaining 'face'.

A second, more granular analysis concerns the delay of sequential alignment. We computed the distribution of all time lags between the withdrawal of a behavior by one participant and the instantiation of that same behavior by another participant. The histogram in figure 3a shows the results collapsed across all coded behaviors. Figures 3b and 3c display the results for gestural behaviors (including gestures, points, nods and shrugs) and for all posture-related behaviors, respectively.

A notable trend is that for all three clusters of behaviors, sequential alignment is very quick: the highest counts of behavioral resonance correspond to a delay of a single second. Further, whereas the sequential alignment rate declines as a function of lag length, it is only at chance level for lags of around 20 seconds (at least when looking at all behaviors; figure 3a).
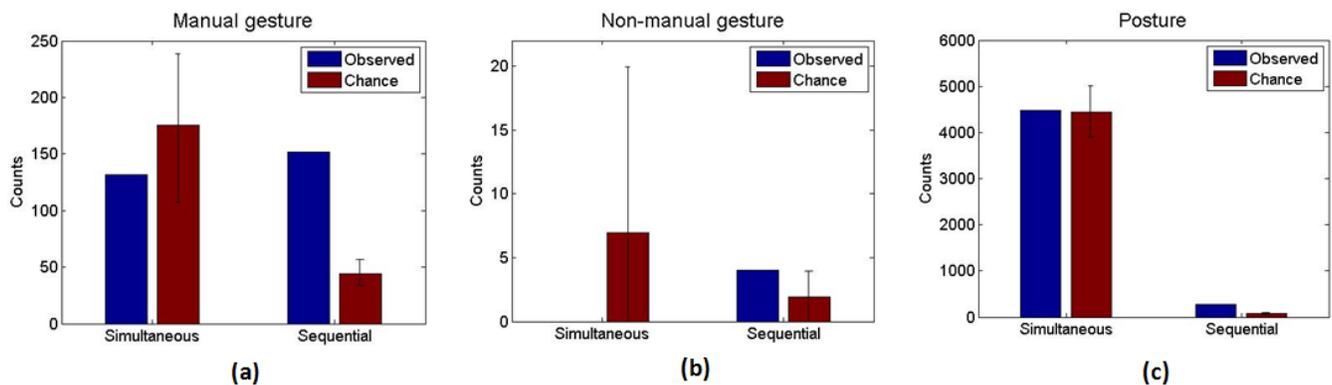

Figure 2: Observed simultaneous and sequential alignment, compared to a random baseline. Error bars represent 95% confidence intervals based on alignment rates in one thousand chance simulations.
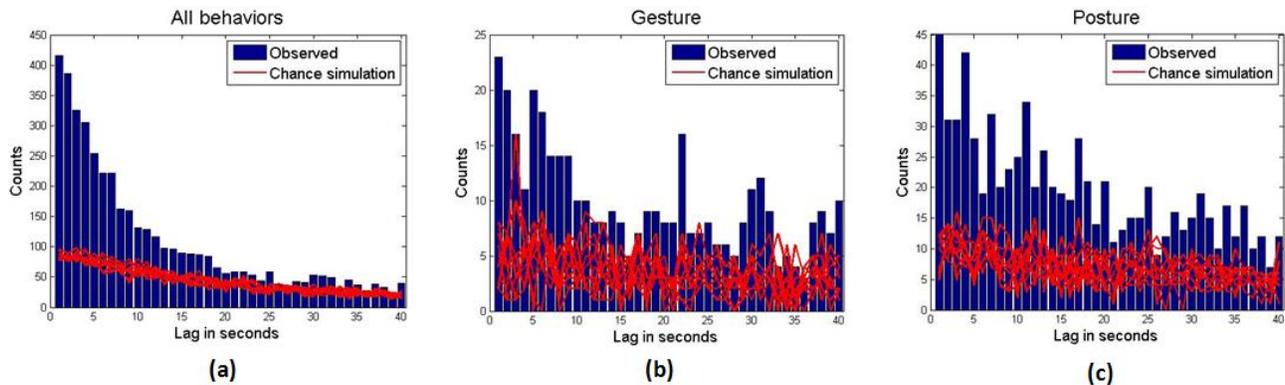
Figure 3: Delay histogram for three clusters of behaviors. Ten out of one thousand chance simulations are plotted.

This suggests that our participants tend to mimic each others' gestural and postural behaviors immediately after observing them, but also with considerable delay.

Conclusions

The research described in this paper has provided insights on the temporal dynamics of behavioral alignment during collaborative remembering. Inspired by the literature and predictions derived from qualitative observations, we tested to what extent people coordinate their behaviors in simultaneous and sequential manners. The results suggest that sequential alignment occurs systematically, whereas simultaneous alignment rates do not exceed chance level. Concurrent performance of the same type of gesture (deictic or non-deictic) by participants in fact happened less often than predicted on the basis of a randomized baseline.

Besides providing interesting results on their own, we believe that combined qualitative-quantitative research methods akin to those described here are a fruitful and welcome addition to experimental research. Especially with regard to cross-behavioral dynamics in joint activities, such methods can contribute to assessing the ecological validity of lab-based experiments and reveal aspects of goal-oriented group interactions that have not been highlighted before.

Acknowledgements

References

Allen, M.L., Haywood, S., Rajendran, T., & Branigan, H.P. (2011). Evidence for syntactic alignment in children with autism. *Developmental Science*, 14, 540–548.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B Resnick, J.M. Levine & S. D. Teasly (Ed.s). *Perspectives on Socially Shared cognition*, Washington: APA Books, 126-149.

de Fornel, M. (1992). The return gesture: Some remarks on context, inference and iconic gesture. In P. Auer & A. di Luzio (eds.), *The Contextualization of Language*. Amsterdam: John Benjamins, 159–176.

Du Bois (2001). Towards a dialogic syntax. Unpublished manuscript.

Furuyama, N. (2000). Gestural interaction between the instructor and the learner in origami instruction. In D. McNeill (Ed.), *Language and gesture*. Cambridge, England: Cambridge University Press, 99–117.

Fusaroli R, Bahrami B, Olsen K, Roepstorff A, Ress G, Frith C, Tylén K (2012). Coming to terms: quantifying the benefits of linguistic coordination. *Psychological Sci.* 23 (8): 931-9.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In B. Ross (Ed.), *The Psychology of Learning and Motivation, 54* (pp. 59-101), Burlington: Academic Press.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. In *Proceedings of Interspeech* 2011. Florence, Italy.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36 (8)*: 1404–1426.

Pickering, M. J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The Art of Conversation Is Coordination Common Ground and the Coupling of Eye Movements During Dialogue. *Psychological science*, *18*(5), 407-413.

Richardson, D, Dale, R. & Shockley, K. (2008) Synchrony and swing in conversation: coordination, temporal dynamics and communication. In I. Wachsmuth, M. Lenzen & G. Knoblich (Eds.), *Embodied Communication in Humans and Machines*. *Oxford:* Oxford University Press.

Shockley K, Baker AA, Richardson MJ, and Fowler CA (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 201–8.