Synchrony and Copying in Conversational Interactions

Kristiina Jokinen
University of Helsinki
Helsinki, Finland
University of Tartu
Tartu, Estonia
`kristiina.jokinen@helsinki.fi`

Siiri Pärkson
University of Tartu
Tartu, Estonia
`siiri.parkson@ut.ee`

Abstract

This paper describes nonverbal communication in conversations, and focuses especially on the interlocutors' synchrony and copying of each other's behaviour. Synchrony and copying indicate the speakers' cooperation with each other, and manifest in the speakers' use of the same words or similar syntactic patterns in their utterances, adjusting their intonation as well as aligning their nonverbal behaviour. We point out some repeated patterns of nonverbal communication in three-party conversations, and offer some interpretations for them.

1 Introduction

One of the fascinating aspects of human conversations is the accurate timing and coordination of the participants' communicative behaviour. Interlocutors react to each others' actions and alternate their turns in a coordinated manner, and they also tend to anticipate and follow the partner's behaviour so that their communication occurs simultaneously and can be described as synchronous activity. This kind of adaptation of the interlocutors to each other's behaviour is often called alignment (Pickering and Garrod, 2004; Katagiri, 2005). Another term that has been used to refer to synchronous behaviour is that of copying or mimicry, which can range from an unintentional copying of a fellow human to an intentional mimic performance. For instance, Caridakis et al. (2007) talk about copying the human behaviour on a virtual character and especially focus on facial expressions and their expressivity, while Mancini et al. (2007) analyze human body movements in order make the virtual character to respond to the user's expressive behaviour appropriately. In virtual agent interactions, mimicry management consists of the subtasks of perception, interpretation, planning, and

animation of the expressions shown by the other person, and it is based on models that represent the user's original expressive behaviour instead of exactly duplicating this.

We can also distinguish synchrony, which functions in a more agent-centred way: although it also requires that the agent has perceived and interpreted the partner's behaviour, it also presupposes that the agent naturally exhibits similar behaviour as the partner: simultaneous behaviour results from the agent's anticipation of the partner's reaction by evaluating the partner's behaviour with respect to the agent's own goals and intentions: synchrony is unconsciously planned rather than intentionally copied from the partner's acting (cf. also Sebanz et al., 2006). The difference between mimicry and synchrony is thus related to the anticipation and coordination of communicative acts: in synchrony, the form of the action originates from the partner's intention to present something in a manner that coincides with the partner's behaviour, while in mimicry only the overt expression of the partner's behaviour is copied.

We have studied synchronous behaviour in three-party conversations and focussed especially on the participants' gestures, body posture, and head movements that occur at the same time. Synchrony can also appear between different communication modalities within a single person, e.g. when one coordinates words with beating gestures, or hand and head movements. However, this kind of intra-partner synchrony is related to the agent's own communication management and has no immediate reference to interaction with the other partner's behaviour, and we will not discuss it here.

In this paper we will focus on inter-partner synchrony, or simultaneous and reciprocal behaviour. Since it signals that the interlocutors are engaged in the interaction and can anticipate the partner's behaviour accurately, we regard syn-

chrony as an indication of the participants' co-operation with each other: the more inclined the participants are to collaborate with their partners, the more synchronous behaviour they show with one another unconsciously. Although the difference between mimicry and synchrony is small on the descriptive level, we aim to distinguish them by referring to intentionality, anticipation and coordination of the speakers' reactions. We say that in mimicry, the speaker synchronizes behaviour in order to produce an affective reaction to the partner's perceived action, but in synchrony, the speaker anticipates a particular behaviour and thus produces spontaneous cooperation, the signal of which is simultaneous similar activity among the partners.

We expect to find a difference concerning the time that it takes for the partner to produce a similar action as the agent, dependent on the time that it takes for the speaker to react. We operationalise the difference by defining the copying behaviour as synchronous activity that has a short time delay with respect to the copied behaviour (due to the time delay in perception, interpretation, planning, and production of an action): the agent copies the partner's gestures, body postures or head movements after a minimum delay of 100ms. It may be difficult to distinguish the two if the delay is a few milliseconds only, and the distinction often depends on the observer's sensitivity to observe the delay too: judgments can vary depending on whether the observer regards the timing of the actions simultaneous or not.

In this paper we describe qualitatively the type of synchrony that occurs between participants, and make a general classification between synchrony and copying by using 100ms as the minimum delay threshold for copying behaviour. We discuss a few examples of synchronous and copying behaviour and try to answer the question if it is possible to distinguish the two in naturally occurring conversations, and if so, which one is more common. In Section 2, we first describe the role and function of gestures and body movement in interactive situations and provide background about the related work. We proceed to describe the data in Section 3, and provide examples of synchrony and copying in Section 4. Finally we discuss some consequences of the work in Section 5, and draw conclusions on the type and function of such behaviour with respect to constructing shared ground in Section 6.

2 Gestures and body movement in interaction

Gestures and body movement have an important role in human communicative behaviour. They are related directly to the information flow of the interactions and they also function in an iconic manner to display the speaker's emotions, attitudes, and mutual relations. They also function on meta-discursive levels (Kendon, 2004; Jokinen and Vanhasalo, 2009), and are used to control and coordinate interaction (Allwood et al., 2007). For example, leaning forward often means interest and leaning backward withdrawing from the conversational situation. Besides displaying the interlocutor's attitudes towards the topic being discussed, body movements can also control interaction by signalling to the partner if they should stop or if they are encouraged to continue further. Such body movements are also used to fill in pauses in conversation: e.g. if the speaker does not want to take the turn, they move backwards. Often the interlocutors also change their position without intending to take the turn in the conversation. They can tacitly state that they are present and have a role in the conversation by adjusting their sitting position appropriately. It is also possible that the body movement is simply related to physical tiredness of staying in a particular position for a long time, but even in this case it can be interpreted as the partner finding the situation uncomfortable and wanting to leave.

Also gaze can control conversations. Gaze signals the speaker's focus of attention and mutual gaze is an important signal in agreeing successful turn-taking (e.g. Jokinen et al., 2009). Gaze may also signal if the speaker wishes to take a turn, or if the turn is offered to another interlocutor (in the latter case, gaze functions in a similar way as pointing, see Fig. 1).

Figure 1 Gaze as a simultaneous pointing device.

In defining communicative gestures and body movement, we follow (Kendon, 2004), who notices that there is a continuum from movements that are perceived as random gesturing to gestures that are understood as communicatively important actions. "Gesture" denotes any possible hand and body movement, but only those which are perceived as communicatively meaningful are *communicative* gestures: potentially all gesturing can be communicatively important if the interlocutors interpret it so. (Sign languages are different in that they form highly structured gesture systems which function by providing abstract representations for communication.)

It is often difficult to assign a clear unambiguous meaning to gesturing and body movements, and often this is not even possible. From the viewpoint of synchronous communication, semantic disambiguation is not necessary since it is not a particular conceptual meaning that is to be conveyed but indication of the partner's collaboration. Any movement can thus function as a starting point for joint gesturing since the partners unconsciously respond to the speaker's gesturing. The speaker also unconsciously reacts to the listener's behavior and would be interesting to study further how the speaker role (the one who speaks) and the contributor role (the one who contributes to conversation) don't necessarily coincide.

If the movements get echoed and amplified by the partner, while the speaker moves back and forth, waves their hands, etc., synchronous behaviour can start. The intuitive nature of such behaviour is often captured in the interlocutors' impressions that it is easy/difficult to talk to the partner: the interlocutors' tacit individual behaviour patterns can either amplify or diminish their joint communicative behaviour, and thus affect their experience of the interaction. To understand what contributes to synchronous behaviour and makes interaction experience pleasant, it is important to investigate how interactions continue and are built up on such movements.

This has an important consequence for synchrony: there are culturally and contextually defined gestures and gesture systems, but only spontaneously elicited gestures that the partner reciprocates can be regarded as truly synchronous in a given situation. Moreover, these gestures can be considered universal in the sense that they are recognized and produced by watching the partner and anticipating their behaviour, without any cultural influence.

Kendon (2004) points out that gestures have a clear peak or stroke, preceeded by a preparation phase, and followed by a post-phase, unlike posture shifts which often are gradual. We define gesture synchrony with respect to the start of the gesture phases: while the length of the speakers' individual preparation phases may vary, it is the timing of the peak that should coincide in their synchronous behaviour.

3 Data

Two videotaped Estonian conversations were used as the basis of our studies. The analysed conversations are altogether about 15 minutes long, and concern three participants talking about plans for a new school building and about inspection of a recently built school building. The situations are role-playing situations, where the participants have adopted the roles of an architect, a school house expert, a town government representative, and a building company representative. The situations are thematically related to each other, i.e. the second conversation is a logical follow-up meeting of the first one, and continuity is supported by two of the participants being assigned the same role in both conversations. Although role-playing may differ from actual situations, it must be noted that people always have a certain role when they are engaged in conversations. Moreover, nonverbal communication and synchrony are mostly unconscious signalling processes, and their conscious modification is not so common; thus nonverbal behaviour may not necessarily differ in role-playing and in spontaneous situations, especially if the participants are familiar with each other as in our case. Since we are not interested in the participants' institutional behaviour, but in their nonverbal communication and synchrony which mostly are unconscious signalling processes, we assume that possible differences between actual and role-playing situations are minimal concerning the purpose of our study.

For the experiment, we manually annotated the behaviour of four individual speakers in the video clips (altogether 15 minutes), and considered dialogue acts, gaze, face, head, turn-taking, feedback and emotion/attitude according to a modified MUMIN scheme (Allwood et al., 2007). Two shorter clips of the same videos were annotated by three annotators and the agreement was measured by Cohen's kappa-coefficient which varied between 40-80% depending on the element. According to the scale proposed by

Rietveld and van Hout (1993), these values correspond to moderate up to substantial agreement. The final annotation is summarized in Table 1 and the relative distribution of different verbal and nonverbal behaviours by the four speakers is shown in Figure 2.

	words	face	gesture	body	all
Sp1	45	213	150	99	462
Sp2	12	75	44	46	165
Sp3	84	242	172	160	574
Sp4	10	127	76	75	278
All	151	657	442	380	1479

Table 1. Statistics of the individual speakers and their behaviour (NV=nonverbal).

There are clear differences between the speakers: Speaker 3 speaks by far the most, and also produces most observable non-verbal communicative acts. Speaker 1 also speaks a lot and is more expressive than Speakers 2 and 4 when it comes to producing facial expressions and hand gestures. The dominance of Speakers 3 and 1 is clearly seen in their gestural and body movements: together they produce more than two thirds of all the observed face, gesture, and body movements. Speakers 2 and 4 speak the least, but differ from each other concerning non-verbal communication: Speaker 2 is the least communicative non-verbally. Synchronous behaviour often occurs between the dominant Speakers 3 and 1, too; this is to be expected as they are the most active in the dialogue.
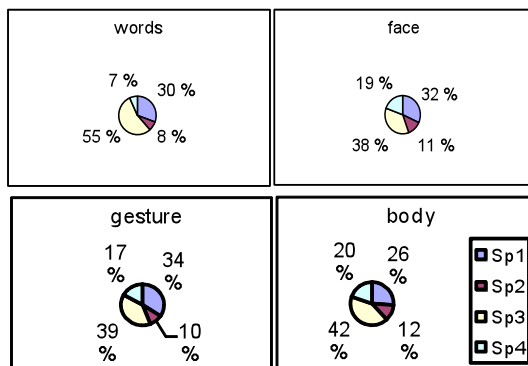

Figure 2. Nonverbal elements in each speaker's behaviour.

When looking at the relative amount of various nonverbal aspects in individual speaker behaviour (Figure 3), we notice that each speaker has majority of the observed nonverbal behaviour encoded in their face and head movement,

supporting the fact that the face is an important means that accompanies speech in a visible and obvious manner. It is interesting that the least talkative participants Speaker 2 and Speaker 4 still have more face and head activity than body or hand movements, but that they use their hands relatively more than their body. This is in accordance with the hypothesis that speech and hand gesturing have an intrinsic connection (see Kendon, 2004), while body movements are not so directly related to speaking.
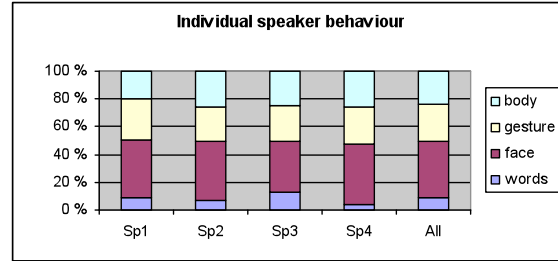

Figure 3. Nonverbal elements in each speaker's behaviour.

4   Synchrony in interaction management

Synchrony and mimicry usually take place between two participants: synchrony that would involve three or more participants seems to be rare, and no such cases appear in our data. In fact, the reason may be obvious as in multi-party dialogues the interlocutors' different roles (speaker, main recipient, side participant) affect their nonverbal behaviour (Battersby, 2011): the interlocutors with different roles react differently to the speaker, and thus it is less probable that their behaviour is synchronised. It must be noted that in two-party dialogues, synchronous situations are not clearly symmetrical either, since one of the partners usually takes the initiative.

Some examples of the synchrony observed in our data are shown in the still-shots in Figures 4-7. Most prominent cases include similar positions with hands crossed (Fig. 4), or hands leaning on the chin (Fig. 5), but also similar body posture (Fig. 6) and the partners' gaze focused on the same object. There are also several examples of beat hand gestures used to emphasize one's arguments (Fig. 7), and the partners copying the behaviour when it is their turn.

Figure 4. Hand-crossing synchrony.

Figure 5. Hand movement synchrony.

Figure 6. Body posture and gaze synchrony.

It is well-known that bursting out to laughing as well as smiling often occur synchronously in smooth conversations (Benus, 2009). Laughing can also create bonds between some participants and leave the others out and thus control the conversation. In our data, for instance, a particular speaker makes a joke which only one of the partners laughs at, and somewhat later, the same speaker makes another joke, which the other partner laughs at. The joking speaker thus seems to control the conversation, and is able to create suitable common grounds so as to engage both partners separately.

Figure 7. Hand beating synchrony.

Similarly to laughing, also nodding often occurs simultaneously, and shows the participants' co-operation and shared understanding. Nodding can also occur as a control signal which directs the participants to talk about certain issues and reach a shared conclusion.

Besides indicating the participants' excite-ment and reinforcing their experience (positive synchrony), synchrony also occurs when a sig-nificant change or communication problem hap-pens in the conversation which the speakers be-come aware of. In order to restore conversational balance and cooperation, the speakers immedi-ately align their behaviours. For example, when a speaker misremembers a fact (last summer was very hot) and the partner hints at misunderstand-ing (children do not go to school in summer), the speaker realises his mistake, and in a moment, a synchrony occurs between the speakers.

Synchronized movements often happen at the start of a new topic and at the change of the speaker. For instance, mutual gaze is an example of this. As one of the participants raises gaze to show that he is ready to take the turn, also the partner simultaneously raises gaze and provides feedback about being interested and listening.

Simultaneous gaze aversion usually also in-dicates the end of a topic or a sequence, and dur-ing the moments of silence, all participants look at their papers or the table. The silence can mark the time the participants need to reflect on the topic, or they simply pretend thinking and hope that someone else will take the turn (this seems to be a steady behaviour pattern especially in the conversation video among the male-only part-ners). However, the breaking of the silence often happens simultaneously.

Speaker's gaze towards the interlocutor can also show that the content of the talk is addressed to the listener or that the listener already has the information (or more information about the is-sue). In our data, the participants do not often look at each other during the discussion, but they

look at the speaker in the beginning and end of the turns, when giving feedback etc. This kind of gazing behaviour may be related to culture-specific conventions.

Finally, it is interesting that, in the conversations, copying of the partner's behaviour is more common than synchrony. Obviously copying of the partner's hand gestures, head movements, and posture shifts helps to create a common ground but it also implies that the participants can easily follow their partner's behaviour and they do this in order to harmonize with their partner. However, it seems less common to get inspired into such simultaneous activity than the synchrony definition presupposes: this would require that the synchronizing partners truly behave in a similar manner as part of their own presentation.

5   Synchrony, copying and cooperation

As mentioned earlier, we consider synchrony as a sign of cooperation: interlocutors cooperate with each other on several levels. In psycholinguistic and social interaction studies such behaviour has been much studied. We base our analysis on the hypothesis that human–human interaction is cooperative activity which emerges from the speakers' capability to act in a relevant and rational manner (Allwood et al., 2007). The basic enablements of communication, Contact, Perception, and Understanding (CPU) must hold for the interaction to proceed smoothly, and consequently, the agents' cooperation can be said to manifest itself to the extent in which the agents can interpret the partner's feedback, and provide relevant feedback on the CPU enablements. Cooperation can manifest itself as a tight collaboration in order to achieve a particular goal, or as similar behaviour patterns that occur when the interlocutors interact and start to align their behaviour with that of the partner. The agents thus constantly monitor themselves (own communication management, see Allwood, 2001) as well as each other, paying attention to the partner's activities and the communicative situation (interaction management), and if any of the enablements is unfulfilled, react to the problems.

In recent years the number of studies concerning synchrony and alignment has increased, maybe due to the new opportunities to experimentally measure and build computational models for simultaneous behaviour. For instance, Benus (2009) studied rhythmic structure of utterances such as pitch accents and syllables with a coupled oscillator model of Wilson and Wilson (2005), and found weak support for the model. They also found that backchannelling had more salient rhythmical characteristics than other turn-taking events.

In general, we can say that synchrony appears between participants who hold together, while asynchronous behaviour is typical between participants who have a contradiction (or pretended contradiction) against each other. The contradiction could be personal or caused by the participants' roles. Synchronous behaviour builds the common ground among the speakers, but we also note that synchrony can also effectively be used to control flow of communication. The speakers can elicit synchronous behaviour e.g. via jokes and nods, and thus express their own individual wishes and viewpoints which, if reinforced through the partner's copying or synchronous behaviour, can further help to achieve the task goals of the interaction itself.

6   Conclusion and future work

We started with the hypotheses that mimicry and synchrony are signs of cooperation through which the participants reinforce their mutual bonds, agreement, and belonging together. According to our analysed data we can confirm this general view: synchrony and mimicry have their own unique role during conversation, and they are signs of the participants' cooperation.

Synchrony may also have other functions. Further analysis with more data is necessary to study these functions in order to produce solid generalizations. It is also important to investigate whether the results hold for other type of conversational activity than the free chatting.

We assume that gesture management deal with the interlocutors' coordinated action of speaking and listening so that only one of the interlocutors speaks at same time. Natural conversations also contain overlaps and silences which can be signals of excitement, cooperation, or ignorance, i.e. they give feedback about the CPU and about the participants' emotional stance. Usually they are short vocalizations as the speakers take their partners cognitive capability into consideration: it is impossible to get one's message across if the speakers speak at the same time.

Relations between interlocutors in interactive situations are usually expressed directly in words but also through nonverbal behaviour. This study focuses on the patterns of nonverbal communica-

tion which can help understand relations between interlocutors, their cooperation, and alignment with each other. Studies on synchrony may be able to explain how the speakers can convey their ideas and viewpoints to their partners, and how they can reach a shared understanding of the communicative situation: by aligning their behaviours, the speakers can experience similar aspects of the situation for which they otherwise have different viewpoints.

References

Jens Allwood. 2001. The Structure of Dialog. In M. Taylor, D. Bouwhuis and F. Nel (Eds.), *The Structure of Multimodal Dialogue II*, 3–24. Amsterdam, Benjamins.

Jens Allwood, Loredana Cerrato, Kristiina Jokinen, Costanza Navarretta and Patrizia Paggio. 2007. The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena. In Martin, et al. (Eds.). Multimodal Corpora for Modelling Human Multimodal Behaviour. *Language Resources and Evaluation*, 41(3-4):273-287.

Stuart Battersby. 2011. Moving Together: the organization of non-verbal cues during multiparty conversation. *Ph.D Thesis*. Queen Mary University of London, UK.

Štefan Benus. 2009. Are We 'in Sync': Turn-taking in Collaborative Dialogues. 10th Annual Conference of the International Speech Communication Association, Interspeech, Special Session: Active Listening & Synchrony. Brighton, UK.

George Caridakis, Amaryllis Raouzaiou, Elisabetta Bevacqua, Maurizio Mancini, Kostas Karpouzis, Lori Malatesta and Catherine Pelachaud. 2007. Virtual Agent Multi-modal Mimicry of Humans. *Language Resources and Evaluation*, 41(3-4):367-388.

Phillip Glenn. 2003. Laughter in Interaction. *Studies in Interactional Sociolinguistics,* volume 18. Cambridge University Press, Cambridge, UK.

Kristiina Jokinen, Masafumi Nishida and Seiichi Yamamoto. 2009. Eye-gaze and Turn-taking. *Proceedings of the third International Universal Communication Symposium, IUCS.* ACM International Conference Proceeding Series. Tokyo, Japan.

Kristiina Jokinen and Minna Vanhasalo. 2009. Stand-up Gestures – Annotation for Communication Management. In C. Navarretta, P. Paggio, J. Allwood, E. Ahlsén and Y. Katagiri. (Eds.). *Proceedings of the NoDaLiDa workshop on Multimodal Communication - from Human Behaviour to Computational Models.* NEALT Proceedings Series, volume 6: 15-20.

Yasuhiro Katagiri. 2005. Interactional alignment in collaborat-ive problem solving dialogues. *Proceedings of the 9th International Pragmatics Conference*. Riva del Garda, Italy.

Adam Kendon. 2004. *Gesture: Visible action as utterance.* New York: Cambridge University Press.

Maurizio Mancini, Ginevra Castellano, Elisabetta Bevacqua and Christopher Peters. 2007. Copying Behaviour of Expressive Motion. *Lecture Notes in Computer Science*, volume 4418. Computer Vision/Computer Graphics Collaboration Techniques. Third International Conference, MIRAGE 2007. Proceedings. 180-191.

Martin J. Pickering and Simon Garrod. 2004. *Towards a Mechanistic Psychology of Dialogue, Behavioral and Brain Sciences*, 27:169-226.

Toni Rietveld and Roeland van Hout. 1993. *Statistical Techniques for the Study of Language and Language Behaviour*. Mouton de Gruyter, Berlin.

Natalie Sebanz, Harold Bekkering and Günther Knoblich. 2006. Joint Action: Bodies and Minds Moving Together. *Trends in Cognitive Science*, 10:70-76.

Margaret Wilson and Thomas P. Wilson. 2005. An Oscillator Model of the Timing of Turn-taking. *Psychonomic Bulletin and Review*, 12(6): 957-968.