Unimodal and Multimodal Co-activation in First Encounters ---- A Case Study

Jens Allwood
SCCIIL, interdisciplinary Center
University of Gothenburg
Göteborg, Sweden
jens@ling.gu.se

Jia Lu
Div. of Communication and Cognition
University of Gothenburg
Göteborg, Sweden
jia.lu@gu.se

Abstract

In human communication, people adapt to each other and jointly activate behavior in different ways. In this pilot study, focusing on one individual (Cf2) in four interactions two types of co-activation, i.e. repetition and reformulation in two modalities, vocal-verbal and gestural are investigated in two Chinese-Chinese and two Chinese-Swedish video-recordings of university students' first encounters. The aim, on the one hand, is to explore features of co-activation that might be specific to Chinese interactions or common to Chinese-Swedish interactions and, on the other hand, to try to see how one person Cf2 adapts to different strangers. In our analysis, we have considered both culture and gender dependent differences. We find that co-activation is more often unimodal than multimodal, and more often involves gesture than speech. We also find that the more similar interlocutors are regarding cultural/ethnic, linguistic, and gender/biological background, the more co-activation takes place, especially in the form of repetition.

Key Words:

Unimodal, multimodal, co-activation, monocultural, intercultural, Chinese, Swedish, vocal-verbal, gestural, culture, gender, interaction

1 Introduction

There are several different approaches to the area of co-activation in communication. One such approach is based on the hypothesis that so called 'mirror neurons' underlie both the production and the perception of movement (Rizzolatti & Arbib, 1998; Arbib, Bonaiuto & Rosta, 2006). Based on neurological studies of 'mirror movement' (Farmer, 2005; Bhattacharya & Lahiri, 2002) and 'mirror neuron' (Gallese & Lakoff, 2005; Arbib, 2005), mechanisms for acting, perceiving, imitation, and pantomime have been identified (Rizzolatti & Arbib, 1998; Ahlsén, 2008). Other theories concerning what we are calling "co-activation" have been labeled 'behavioral adaptation' (Galegher & Kraut, 1992), 'adaptive response' (Buck, 1984; Burgoon, Stern & Dillman, 1995; Cappella, 1991), 'imitation' (Ahlsén, 2008; Arbib, 2005), bodily coordination (Ivry & Richardson, 2002; Semjen & Ivry, 2001), 'alignment and automatized coordination' (Pickering & Garrod, 2004), and the phenomena considered are usually regarded by the cited authors as a basic and crucial part of human communication and language development. The terms chosen in the mentioned approaches all point to different but probably related aspects of 'bodily coordination'. In this study, we use the term 'co-activation' to refer to the occurrence of similar vocal-verbal and gestural behaviors that occur in different communicators either sequentially or simultaneously, in order to serve the purpose of coordinating human communication. We use the term "gestural" for all visible communicative body movements and the term "vocal-verbal" to distinguish verbal expressions that are vocal from verbal expressions that are gestural, e.g. the gestural words of deaf sign language or the head nods and head shakes used in feedback which we also regard as gestural words.

2 Types of Co-activation

We will take both vocal-verbal and gestural co-activation into account. An interesting part of the relevant behavior consists of communicative feedback (cf. Allwood, Ahlsén & Nivre, 1992; Allwood & Cerrato, 2003; Grammer, Allwood,

Ahlsén & Kopp, 2008). Co-activation can occur vocally through words or phrases, some of which consist of repetitions or reformulations, e.g. B says 'that's all right' after A says 'that's all right' (repetition), or B says 'that's fine' after A says 'that's all right' (reformulation). Co-activation can also occur through gestures; we have coded head movements (down-nod, up-nod, and shake), facial expressions (eyebrow frown, eyebrow rise, gaze up, gaze down, gaze at the other interlocutor, gaze sideways i.e. gaze left or right, smile, scowl (mouth open in a circle, and mouth corners down), posture shifts, shoulder movements (mainly shoulder shrugs), and hand movements as well as through combinations of vocal and gestural behavior, i.e. laughter, chuckle (basically a smile plus a laughing sound with a low pitch and intensity) or giggle (a smile plus a laughing sound with a high pitch and intensity, which are repeated or reformulated, e.g. B smiles after A smiles (repetition), or B chuckles to express friendliness after A has smiled in a friendly way (reformulation). The idea is that a gestural repetition involves use of "the same gesture" in terms of both function and expression, while a gestural reformulation also often involves use of a "similar gesture" and a "similar function". However, the requirement on similarity in function is stronger than the requirement on similarity in expression since, for instance, a negative headshake can be reformulated as a negative hand movement. We admit that as far as reformulations go, the boundaries concerning what is to be regarded as "similar" are somewhat vague both with regard to vocal and gestural expressions and their functions. Operationally, we have tried to restrict what is regarded as similar fairly narrowly to units that serve the same function in a fairly clear sense.

Below, we will use the term "unimodal" for co-activation that is vocal-verbal (only) or gestural (only) and "multimodal" for co-activation that is vocal-verbal plus gestural. In this paper, we restrict our study of co-activation to repetitions and reformulations, while not denying that the concept of co-activation has a wider application.

3. Purpose

This paper primarily investigates three questions. First, what vocal-verbal and gestural behaviors occur in unimodal and multimodal co-activation? Second, are different types of co-activation used in mono-cultural and intercultural interactions? Third, are there any gender differences?

4. Data and Method

The study is based on four video-recordings of face-to-face dyadic dialogs between Chinese and Swedish university students. In order to make a pilot case study of co-activation with respect to differences in culture and gender, one Chinese female subject (Cf2) was studied both in two Chinese-Chinese and two Chinese-Swedish dialogs that varied in the gender of her interlocutors (see Table 1). This allows us to see how the gender of a communicative partner might influence one and the same person (Cf2). Thus, in the mono-cultural interactions, Cf2 was studied with a Chinese female (Cf1) and a Chinese male (Cm1) and in the intercultural interactions, Cf2 was studied with a Swedish female (Sf2) and a Swedish male (Sm2). Since the number of examined recordings is small, a more representative study will require more data.

Recording	*Participants*	*Time Length*	*Language*
Dial.1	Cf2--Cf1	7:00 min.	Chinese
Dial.2	Cf2--Cm1	7:00 min.	Chinese
Dial.3	Cf2--Sf2	7:00 min.	English
Dial.4	Cf2--Sm2	7:00 min.	English

Table 1: The studied video-recordings ***(Note:*** *C=Chinese, S=Swedish, f=female, and m=male.)*

Our study is focused on how strangers who have no earlier acquaintance go about the task of getting to know each other. Each interaction was video-filmed by three video cameras (left-, center-, and right-position) with each interlocutor in a standing position (see Figure 1). The main subject Cf2 was video-recorded four times, and her counterparts Cf1, Cm1, Sf2 and Sm2 were video-recorded once each to provide different adaptation contexts for Cf2. Each video recording lasted approximately seven to ten minutes, but only the first seven minutes were analyzed in detail in the present study.

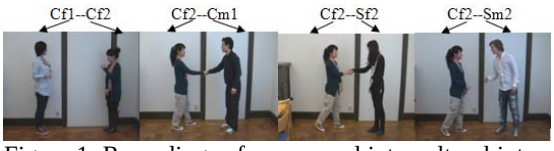

Figure 1: Recordings of mono- and intercultural interactions

The video-recorded data was transcribed and checked according to the GTS (Göteborg Transcription Standard) version 6.2 (Nivre, 1999). To increase reliability, each video recording has one transcriber and two independent checkers. All the video-recordings were manually annotated following the MUMIN multimodal coding scheme (Allwood, Cerrato, Jokinen, Navarretta & Paggio, 2007).

5. Analysis and Results

Below we will now analyze the four recorded dialogs from the perspective of whether the co-activation occurring is multimodal or unimodal.

5.1 Overview

Results concerning co-activation through repetition and reformulation, for all five participants, are presented in Tables 2 and 3. Table 2 shows that there is more unimodal gestural than unimodal vocal-verbal co-activation (171-69), while in contrast, there are only 19 cases of multimodal co-activation, for all participants in the four recordings.

Modality	***Type***	***Total***
Vocal-verbal Unimodal	Repetition	57
	Reformulation	12
	Total	**69**
Gestural Unimodal	Repetition	111
	Reformulation	60
	Total	**171**
Vocal-verbal + Gestural Multimodal	Repetition	6
	Reformulation	13
	Total	**19**

Table 2: Total number of unimodal and multi-modal co-activations (including both Chinese and Swedish participants)

Modality	***Type***	***Mon.***	***Int.***	***Total***
Vocal-verbal (only)	Rep.	12	11	23
	Ref.	2	0	2
	Total	**14**	**11**	**25**
Gestural (only)	Rep.	31	34	65
	Ref.	14	15	29
	Total	**45**	**49**	**94**
Vocal-verbal + Gestural	Rep.	3	0	3
	Ref.	3	1	4
	Total	**6**	**1**	**7**

Table 3: Cf2's unimodal and multi-modal co-activation *(Mon.=mono-cultural, Int.=intercultural, Ref.=reformulation, Rep.=repetition)*

In addition, we can see (Table 3) that the main subject Cf2 exhibits the same proportions between vocal-verbal and gestural and multimodal co-activation as those observed for the group as a whole (Table 2), but that the differences between Cf2's behavior in the mono-cultural and intercultural situation, are too small to be significant.

5.2 Unimodal Co-activation

In this section, unimodal co-activation i.e. vocal-verbal (vocal-verbal only) and gestural (gestural only) co-activation is studied more in detail.

5.2.1 Unimodal Vocal-verbal Co-activation

Below we will exemplify unimodal vocal-verbal co-activation as it can be observed through repetitions and reformulations. Excerpt 1 shows how the vocal-verbal expression 'wang you' ('turn to the right' in English) is repeated by speaker Cf2, while Excerpt 2 shows how 'hello' is reformulated to 'hi' by speaker Cf2.

Excerpt[1] 1 vocal-verbal unimodal repetition:	
Original Transcription	*Literal English Trans.*
$Cf1: <1 en >1 /// <2 wo men shi wang zuo >2 /// ni men shi **wang you** …	$Cf1: <1 yeah >1 /// <2 we turn to the left >2 /// you **turn to the right** …
@ <1 VFB; CPU confirmation >1 @ <2 VFB; CPU confirmation >2…	
$Cf2: <1 a /// dui dui dui >1 <2 **wang you** >2 …	$Cf2: <1 ah /// right right right >1 <2 **turn to the right** >2 …
@ <1 VFB; CPU confirmation >1 @ <2 VFB; CPU confirmation >2…	
Excerpt 2 vocal-verbal unimodal reformulation:	
$Sf2: **hello** $Cf2: **hi** < | > e1 @ < general face: giggle >, < hand start: Sf2, Cf2 shake hands >	

The vocal-verbal unimodal co-activations can be classified in terms of phrase categories and parts of speech. In Excerpt 1, 'wang you' ('turn to the right' in English) is a verb phrase that is repeated as feedback; in Excerpt 2, 'hello' and 'hi' are both interjections.

[1] The excerpts in this paper are extracted from transcriptions of the studied recordings. In GTS, $ identifies a speaker. Angular brackets < > indicate the scope of a comment, and the number identifies a corresponding comment. The symbol @ initiates the corresponding comment. The number of slashes (/, //, ///) indicate the length of a pause. Curled brackets { } contains letters of a written word form that were not pronounced in the spoken form. < | > indicates that a gesture without vocal-verbal information is inserted in a pause. In our coding, VFB= vocal-verbal feedback, GFB= gestural feedback, CPUE/A= contact, perception, understanding, emotion/attitude.

Feature	Frequency	Examples of repeated expressions
N/NP	37 (65%)	Hobbies; The American idol
V/VP	9 (16%)	Yao qiu *'require';* Hai pa jin qin *'(be) afraid of intermarriage'*
Adj	3 (5%)	Similar
Sentence	2 (4%)	Vad sa du *'what did you say'*
Int	2 (4%)	Hej *'hi'*
Adv	2 (4%)	Just
Pron	1 (1%)	Ta-men *'they'*
Prep	1 (1%)	(Shi) zai *'(be) at'*
Total	57 (100%)	
Relation to FB: 34 repetitions, 60%, are feedback		

Table 4: Grammatical categories of all vocal-verbal unimodal repetitions (The intercultural dialogs, although mainly in English, include a few Swedish expressions)

Table 4 shows the grammatical categories of the unimodal vocal-verbal repetitions; N (noun) and NP (noun phrase) (65%), V (verb) and VP (verb phrase) (16%). We may note that 60% of all the unimodal vocal-verbal repetitions have a feedback function, which indicates that co-activation and feedback are closely connected.

Feature	Frequency	Example
N/NP	5 (42%)	Bei jing *'backgound'* → Gong zuo bei jing *'working background'*
Adj	3 (25%)	Ting hao de *'(it is) very good'* → Bu cuo *'not wrong'*
V/VP	2 (17%)	Guo guo *'pass pass'* → Pass (English)
Pronoun	1 (8%)	I saw it → You saw it.
V/Prep	1 (1%)	Wang you *'(turn) to the right'* → (zai) you bian *'on the right'*
Total	12 (100%)	
Relation to FB: 3 reformulations, 25%, are FB		

Table 5: Grammatical categories of all unimodal vocal-verbal reformulations

Concerning unimodal vocal-verbal reformulations, the most common types are N/NP (42%), Adj (adjective) (25%), and V/VP (17%) (cf. Table 5). 25% of the vocal-verbal reformulations have a feedback function, which again, although weaker than for repetition, shows a link between co-activation and feedback.

We have seen in Table 2 (see also Table 6 below), that there are 57 repetitions and 12 unimodal vocal-verbal reformulations, altogether 69 unimodal vocal-verbal instances of co-activation (produced by both Chinese and Swedish paticipants). Thus, the number of vocal-verbal unimodal repetitions is approximately five times as large as that of vocal-verbal unimodal reformulations.

Vocal-verbal unimodal	*Dial.1*		*Dial.2*		*Dial.3*		*Dial.4*		*Total*
	Cf1	*Cf2*	*Cm1*	*Cf2*	*Sf2*	*Cf2*	*Sm2*	*Cf2*	
Repetition	9	7	10	5	5	3	10	8	57
Reformulation	2	0	3	2	2	0	3	0	12
Total	11	7	13	7	7	3	13	8	69

Table 6: Vocal-verbal unimodal co-activation in the recordings

We have chosen to study the Chinese subject Cf2, varying the gender and/or culture of her interlocutor. Cf2 shows the same tendency as the group as a whole using more unimodal (23) vocal-verbal repetitions than reformulations (2), as can be seen from Table 6. She used roughly the same number of unimodal vocal-verbal repetitions and reformulations in the Chinese mono-cultural interactions (12 (i.e. 7+5) and 2 (i.e. 0+2)) as in the intercultural interactions with the Swedes (11 (i.e. 3+8) and 0 (i.e. 0+0)).

With respect to the gender differences in using unimodal vocal-verbal co-activation, Cf2's interactions are illustrative. As shown in Table 6, Cf2 had slightly more vocal-verbal unimodal co-activation with males (Cm1(13) + Sm2(13)) than with females (Cf1(11) + Sf2 (7)). The number of cases is too small to allow any claim about gender difference in Cf2's interactions with Chinese interlocutors.

Vocal-verbal unimodal	***Dial.1 with Cf1***	***Dial.2 with Cm1***	***Dial.3 with Sf2***	***Dial.4 with Sm2***	***Total***
Repetition	7	5	3	8	23
Reformulation	0	2	0	0	2
Total	7	7	3	8	25

Table 7: Cf2's unimodal vocal-verbal co-activation

However, turning to repetitions and reformulations, in Dialogs 3 and 4 (see Table 7), Cf2 used more unimodal vocal-verbal repetitions with the Swedish male (8) than with the Swedish female (3) and Cf2 did not use any unimodal vocal-verbal reformulations with Swedish interlocutors.

5.2.2 Unimodal Gestural Co-activation

We have found totally 171 instances of unimodal gestural co-activation in all four analyzed dialogs. Of these 111 were repetitions and 60 reformulations (see Table 8).

Gestural unimodal	*Dial.1*		*Dial.2*		*Dial.3*		*Dial.4*		*Total*
	Cf1	*Cf2*	*Cm1*	*Cf2*	*Sf2*	*Cf2*	*Sm2*	*Cf2*	
Repetition	7	20	13	11	13	23	13	11	111
Reformulation	10	5	7	9	8	7	6	8	60
Total	17	25	20	20	21	30	19	19	171

Table 8: Unimodal gestural co-activation in the recordings

Thus, the number of unimodal gestural repetitions is approximately twice as many as that of unimodal gestural reformulations.

Excerpt 3 gestural unimodal smile repetition:	
Original Transcription	*Literal English Trans.*
$Cf2: <1 en /// >1 <2 | >2	$Cf2: <1yeah///>1 <2 |>2
@ <1 VFB; CPU confirmation >1, <1 GFB head: nods; CPU confirmation >1 @ <2GFB general face:**smile**;CPUE/A friendliness>2	
$Cm1: <1 | >1 <2 ou >2 <3 wo shi >3 <4 wo shi <5 hui zu >5 >4	$Cm1: <1|>1<2 oh >2 <3 i am >3 <4 i am (from) <5 hui nationality >5 >4
@ <1 GFB general face: **smile**; CPUE/A surprise/happiness >1 @ <2 VFB; CPU >2…	
Excerpt 4 gestural unimodal reformulation:	
$Cf2: [2 <1 oh >1 <2 yeah similar >2]2 // [3 in the]3 pronunciation [4 <3 // >3]4 // and … @ <3 general face: **giggle** >3 $Sf2: [3 <1 yeah >1 <2 | >2]3 @ <1 VFB; CPUE/A agreement >1, <1 GFB head: nods; CPUE/A agreement R >1 @ <2 GFB general face: **chuckle**; CPUE/A friendliness >2	

Excerpt 3, above shows how a smile is repeated unimodally by Cm1, and Excerpt 4 how Cf2's giggle is reformulated unimodally into a chuckle by Sf2. The unimodal gestural co-activations in Excerpts 3 and 4 are both related to the behavioral group smile/ giggle/ laughter/chuckle which often express friendliness, surprise or happiness, all of which are expectable and fairly common in first acquaintance dialogs.

In general, we have found (see Table 9, below) that unimodal gestural repetitions most frequently involve the following body parts; head (50%), general face (especially smile/ giggle/chuckle/ laughter) (37%), and gaze (6%), and that 69% of the unimodal gestural repetitions have a feedback (FB) function.

Co-activated gestures	**Freq.**	**Example**
Head (nod/ up-nod/ shake/ tilt/ others)	55 (50%)	$Cf2: <1 i'm li yun / <2 nice to meet >2 you >1 … @ <1 hand: **Cf2, Sm2 shake hands** >1 @ <2 GFB head: **Sm2 nod**; CPU >2, <2 head: **nod** >2 $Sm2:…<2i'm jesper>2 @ <2head: **Cf2 nods**>2 $Cf2: < oh > @ < VFB; CPU >, < GFB head: **nod**; CPU >
General face (smile/ giggle/chuckle/laughter)	41 (37%)	
Gaze (up/ down/ side-ways/ around)	7 (6%)	
Posture movement	4 (4%)	
Hand movement	3 (3%)	
Arm movement	0 (0%)	
Total	110(100%)	
Relation to FB: 76 (69%), have a feedback function		

Table 9: Body parts involved in gestural repetition

In Table 10 below, we can see the corresponding figures for gestural reformulation.

Co-activated gestures	**Frequency**	**Example**
General face (smile/ giggle/ chuckle/ laughter)	77 (62%)	$J: <1 yeah >1 it's kin+ i wou{ld} think it's kind of hard for you to <2 understand swedish [49 // >2 <3 elle{r}]49 sevenska >3 @ <1 VFB; CPUE/A agreement >1, <1 GFB **gaze: down**; CPUE/A hesitation O >1 … $L: [49 < (...) >]49 @ < **gaze around** >
Head (nod/up-nod/ shake/ tilt/ others)	17 (14%)	
Gaze (up/ down/ side-ways/ around)	13 (10%)	
Hand movement	8 (6%)	
Posture movement	8 (6%)	
Arm movement	2 (2%)	
Total	125(100%)	
Relation to FB: 71 raw frequencies, 57%, are FB		

Table 10: Body parts involved in unimodal gestural reformulation

Unimodal gestural reformulation is most frequently facial (especially smile/ giggle/ chuckle/ laughter) (62%), head (14%), and gaze movement (10%) (see Table 10), and 57% of the unimodal gestural reformulations have a feedback (FB) function.

Gestural unimodal	***Dial.1 with Cf1***	***Dial.2 with Cm1***	***Dial.3 with Sf2***	***Dial.4 with Sm2***	***Total***
Repetition	20	11	23	11	65
Reformulation	5	9	7	8	29
Total	25	20	30	19	94

Table 11: Cf2`s unimodal gestural co-activation

Turning back to Cf2, Table 11, above, shows that she used more than twice as many unimodal gestural repetitions (65) as reformulations (29). She further used almost the same number of unimodal gestural repetitions and reformulations with Chinese as with Swedish interlocutors: Repetitions; Chinese 31 (i.e. 20+11)) and Swedes

34 (i.e. 23+11); Reformulations; Chinese 14 (i.e. 5+9) and Swedes 15 reformulations (i.e. 7+8).

Concerning gender differences, Cf2 used roughly twice as many repetitive gestures when she interacts with females (43) as with males (22), irrespective of culture (cf. Table 11) and she used slightly more unimodal gestural reformulations with males than with females (as 9 to 5 in monocultural dialogs, and 8 to 7 in intercultural dialogs). That is, in both mono-cultural and intercultural interactions, Cf2 had more unimodal gestural repetitions with females and slightly more unimodal gestural reformulations with males.

5.3 Multimodal Co-activation

We now turn to multimodal co-activation. As can be seen from Table 12, there are totally 19 instances of multimodal co-activation, including both Chinese and Swedish subjects.

Multimodal V+G	*Dial.1*		*Dial.2*		*Dial.3*		*Dial.4*		*Total*
	Cf1	*Cf2*	*Cm1*	*Cf2*	*Sf2*	*Cf2*	*Sm2*	*Cf2*	
Repetition	0	2	1	1	1	0	1	0	6
Reformulation	1	0	1	3	1	1	6	0	13
Total	1	2	2	4	2	1	7	0	19

Table 12: Multimodal co-activation (*V+G=vocal-verbal+gestural)*

Of these, 6 are multimodal repetitions (see Excerpt 5) and 13 reformulations (see Excerpt 6, below). Thus, the number of multimodal reformulations is approximately twice as many as that of the multimodal repetitions.

Excerpt 5 multimodal repetition:	
$Sm2: we <1 call it >1 <2 **peking** >2 @ <1 general face: Cf2 chuckle >1 @ <2 name: city >2, <2 **smile** >2 $Cf2: <1 | >1 <2 yeah >2 <3 **peking** >3 [5 //]5 <4 en >4 // and u1… @ <3 VFB; CPU confirmation >3, <3 GFB general face: **smile**; CPUE/A friendliness O >3, <3 name: city >3	
Excerpt 6 multimodal reformulation:	
Original Transcription	*Literal English Translation*
$Cm1: < **hai** >	$Cm1: < **hi** >
@ < **right hand shake** >, < **smile** >	
$Cf2: < **hai ni hao** >	$Cf2: < **hi hello** >
@ < **right hand shake** >, < **smile** >	

In Excerpt 5, the multimodal unit, 'peking' + a smile, is repeated by speaker Cf2. In Excerpt 6, the multimodal unit 'hai' ('hi' in English) plus handshake and smile, is reformulated by speaker Cf2 into 'hai ni hao' ('hi/ hello' in English) plus a handshake and smile.

Returning to Cf2, she did not repeat or reformulate multi-modally very often in either monocultural or intercultural interactions. In both types of dialog, she had a similar number of multimodal reformulations (4) and multimodal repetitions (3). See Table 13, below.

Multimodal V+G	***Dial.1 with Cf1***	***Dial.2 with Cm1***	***Dial.3 with Sf2***	***Dial.4 with Sm2***	***Total***
Repetition	2	1	0	0	3
Reformulation	0	3	1	0	4
Total	2	4	1	0	7

Table 13: Dynamic features of multimodal co-activation made by Cf2

She used slightly more multimodal repetitions and reformulations with the Chinese (6) than with the Swedish (1) interlocutors: Repetitions; 3 (i.e. 2+1) versus 0 (i.e. 0+0) and Reformulations; 3 (i.e. 0+3) versus 1 (i.e. 1+0). That is, Cf2 used slightly more multimodal co-activation in monocultural interactions (6) than in intercultural interactions (1).

With respect to the possible influence of gender, when interacting with Cf2, males used more multimodal co-activation than females (Cm1 had 2 and Cf1 had 1; Sm2 had 7 and Sf2 had 2). Cf2 used roughly the same number of multimodal repetitions with the Chinese female (2) and the Chinese male (1); however, she used slightly more multimodal reformulations with the Chinese male (3) than with the Chinese female (0). In the intercultural interactions, Cf2 used roughly the same number multimodal reformulations with the Swedish female (with a frequency of 1) as with the Swedish male (0). Cf2 did not use any multimodal repetitions with the Swedish interlocutors at all.

6. Discussion

In section 5, we have found more unimodal co-activation instances than multimodal ones (approximately 12 times as many) in the examined recordings. Possibly this indicates that co-activation in human communication is more unimodal than multimodal. We also found that unimodal gestural co-activation was twice as common as unimodal vocal-verbal co-activation. This possibly shows that co-activation in human communication is more dependent on gestures than on speech. In addition, we found that multimodality plays a relatively less important role

than unimodality for co-activation in the first encounters we have studied.

Both Chinese and Swedish participants used more unimodal vocal-verbal and gestural repetitions than unimodal reformulations in their co-activation. This may be an automatic effect of 'mirror neurons', or because in first encounters interlocutors repeat each other's vocal-verbal information, in order to confirm whether they have perceived and understood the information correctly. Both Chinese and Swedish subjects used more multimodal reformulations than multimodal repetitions, possibly because it is more difficult to repeat complex multimodal units of behavior. Unimodal behavior may be easier to repeat, especially vocal-verbal unimodal behavior; whereas, multimodal behavior is more difficult to repeat but easier to reformulate.

We found that both vocal-verbal and gestural unimodal co-activation occurred more frequently with the males than with the females when they were interacting with the Chinese female Cf2, in both mono-cultural and intercultural interactions. Specifically, we found that the males used more unimodal gestural repetition than the females, when interacting with Cf2. Possibly, this is because males are less socially elaborating than females, repeating more and reformulating less.

We have also observed what parts of speech or what parts of the body were involved in unimodal vocal-verbal or gestural co-activation. We found that nouns or noun phrases and verbs or verb phrases comprise most of the unimodal vocal-verbal co-activation, and that more than half of them have a feedback function. Possibly this is because nouns and verbs mostly provide the core of the topic being talked about, and feedback is needed for managing and keeping the interaction going. Further, we found that head, general face (especially smile, chuckle, giggle, laughter), and gaze movements are the most common unimodally co-activated gestures. This may be, because head and face are central in human interaction, so that people attend and react more to the information carried by head movements and facial expressions. For instance, they often try to be friendly in a first encounter and therefore smile or laugh, or they express emotional rapport, hesitation/uncertainty, and/ or interest through gaze movement. Again, more than 50% of the unimodal gestural co-activation has a feedback function, which indicates that giving and eliciting feedback plays a very important role in co-activation in human communication.

If we turn to features that might be specifically Chinese, Cf2 exhibited slightly more vocal-verbal and multimodal co-activation in the mono-cultural interactions than in the intercultural interactions, but more unimodal gestural co-activation in the intercultural ones (cf. table 3, above). The reason for this might be that she felt more comfortable with the other person's vocal-verbal behavior when both of them come from the same cultural and linguistic background, not least for reasons of automatic linguistic proficiency. Perhaps this makes vocal-verbal co-activation easier in mono-cultural interactions, and gestural co-activation, relatively speaking, more comfortable in intercultural interactions.

Cf2 used more unimodal gestural repetition with the same gender and more unimodal gestural reformulation with the other gender in both mono-cultural and intercultural interactions. The reason could be that it is easier to repeat gestural behavior from persons of the same gender. It may be that the more similarities interlocutors share in cultural and biological background, the more repetitions they produce.

7. Limitation of research

Our study has some limitations. First of all, since there are only two Chinese-Chinese mono-cultural and two Chinese-Swedish intercultural interactions, involving two Chinese females, one Chinese male, one Swedish female and one Swedish male, the preliminary results and conclusions are all very tentative.

Second, the results based on Cf2 may be dependent on Cf2 as an individual, and other results may be activity dependent. This necessitates further studies in the future.

Third, Cf2 was video-recorded four times. This means that Cf2 had more experience in the later recordings, and to some extent she was used to communicating with a stranger before a video camera.

Fourth, this pilot study focuses on a small number of Chinese overseas and Swedish native university students in first encounters. So it is unclear to what extent it can be regarded as repre-

senting the general Chinese features of unimodal and multimodal co-activation.

8. Conclusions

The aim of this study was to explore the following research questions: What are the features of co-activation with strangers in vocal-verbal and gestural behavior? Do interlocutors use different types of co-activation in mono-cultural and intercultural interactions? Are there any gender influences?

Because our study is small in size, below are only some suggestions and tendencies that can be seen in our data. Concerning the Chinese female participant Cf2's co-activation in mono-cultural and intercultural interactions, she had slightly more unimodal vocal-verbal and multimodal co-activation in mono-cultural than in intercultural interactions but for unimodal gestural co-activation the difference went in the other direction and since the differences, in any case, were too small to be significant, we do not really have an answer to the question of whether interlocutors use different types of co-activation in mono-cultural and intercultural interactions.

Second, Cf2 used more unimodal gestural repetitions with the same gender in both mono-cultural and intercultural interactions. She also used more multimodal repetitions with the same gender in mono-cultural interactions. This suggests that it is easier for an interlocutor to repeat gestural unimodal and multimodal behaviors when the gender of the interlocutors is the same, possibly for biological reasons. It also supports the view that the more similarities interlocutors share in cultural/ethnic, linguistic, and gender/biological background, the more co-activation is possible.

We also found some common trends for Chinese and Swedish interlocutors. First, unimodal gestural co-activation was more common than unimodal vocal-verbal co-activation, which points to easier access to gestures than to speech or to a greater role for the visual modality than for the auditory modality in co-activation. Multimodality, thus, seems to play a relatively less important role in co-activation, at least in the first encounters we have studied. Second, both Chinese and Swedish interlocutors used more unimodal vocal-verbal and gestural repetitions than unimodal reformulations, but they used more multimodal reformulations than multimodal repetitions. Some possible explanations for this could be that they are making a conscious effort at giving vocal-verbal confirmatory feedback on perception and understanding, or that they are reacting as a result of unconscious mechanical effects of 'mirror neurons'. Another possibility is that it is more difficult to repeat multimodal unit of behaviors, at least in a first encounter. These all necessitate further study.

It was also found that nouns, verbs, and feedback expressions comprised most of the vocal-verbal unimodal co-activation; head, general face (especially smile, chuckle, giggle, laughter), and gaze were the most common unimodally co-activated gestures. This may be because nouns and verbs often are centrally related to the topic, and feedback is used for managing interaction; head and face attract more attention in human interactions, and interlocutors try to be friendly in first encounters or express emotional rapport, hesitation/ uncertainty, and/ or interest through gaze movement.

Males used more vocal-verbal unimodal co-activation and more gestural unimodal repetition but less gestural unimodal reformulation than females in both mono-cultural and intercultural interactions. We speculate that the reason for this might be that males are less socially elaborating than females.

Since our data and activity variation are quite limited, further research is needed to attempt generalizations about cultural and gender differences. This pilot study can therefore mostly contribute to a general description of how people adapt to others through co-activation of vocal-verbal and gestural unimodal and multimodal behavior.

Acknowledgement:

We express our gratitude for support to the VR (Swedish Research Council) project "Återkopplingsprocesser" and to the NOS-HS (Nordic Research Council) project NOMCO. We would also like to thank Elisabeth Ahlsén for discussions and comments and Alexander Holender and Yansi Xu for reliability checking work done at the SCCIL Interdisciplinary Research Center at the university of Gothenburg (Sweden). Finally, we thank our reviewers for constructive comments.

References

Allwood, J., Ahlsén, E., Nivre, J. (1992). On the Semantics and Pragmatics of Linguistic Feedback. *Journal of Semantics* 9, 1-26.

Allwood, J. & Cerrato, L. (2003). A Study of Gestural Feedback Expressions. In P. Paggio, K. Jokinen & A. Jönsson (eds) *First Nordic Symposium on Multimodal Communication.* ISSN 1600-339X, 7-22. Copenhagen.

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. & Paggio, P. (2007). The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing. In J. C. Martin et al. (eds) *Multimodal Corpora for Modelling Human Multimodal Behaviour*. Special issue of the International Journal of Language Resources and Evaluation. Springer. Vol.41, no.3-4, pp.273–287.

Ahlsén, E. (2008). Embodiment in communication-aphasia, apraxia and the possible role of mirroring and imitation. *Clinical Linguistics and Phonetics*, April-May 2008; 22 (4-5): 311-315.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Brain and Behavioral Sciences*, 28, 105–124.

Arbib, M. A., Bonaiuto, J. & Rosta, E. (2006) The mirror system hypothesis: From a macaque-like mirror system to imitation. *In Proceedings of the 6th International Conference on the Evolution of Language*, 3--10.

Bhattacharya, A. & Lahiri, A. (2002). Mirror Movement in Clinical Practice. *Indian Academy of Clinical Medicine.* Vol. 3, No. 2, 177-81.

Buck, R. (1984). *The communication of emotion*. New York: Guilford Press.

Burgoon, J. K., Stern, L. A. & Dillman, L. (1995). *Interpersonal adaptation: Dyadic interaction patterns*. Cambridge, UK: Cambridge University Press.

Cappella, J. N. (1991). The biological origins of automated patterns of human interaction. *Communication Theory*, 1, 4-35.

Farmer, S. F. (2005). Mirror movements in neurology. *Neurol Neurosurg Psychiatry* 76, 1330. Online ISSN 1468-330X.

Galegher, J. & Kraut, R. E. (1992). Computer-mediated communication and collaborative writing: media influence and adaptation to communication constraints. *CSCW '92 Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, ISBN:0-89791-542-9, 155-162.

Gallese, V. & Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22 (3/4), 455–479.

Grammer, K., Allwood, J., Ahlsén, E., Kopp, S. (2008). A Framework for Analyzing Embodied Communicative Feedback in Multimodal Corpora. *JLRE (Special Issue on Multimodal Corpora)*. J.C. Martin (ed.) No. 66466.

Ivry, R. B. & Richardson, R. E. (2002). Temporal control and coordination: The multiple timer model. *Brain and Cognition*, 48, 117–132.

Kolb, B. & Whishaw, I. Q. (2003). *Fundamentals of Human Neuropsychology*. New York: Freeman.

Nivre, J. (1999) *Göteborg Transcription Standard. Version 6*.2, pp. 38. Göteborg: Göteborg University, Department of Linguistics.

Pickering, M. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169-225.

Rizzolatti, G. & Arbib M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21,188-194.

Semjen, A., & Ivry, R. B. (2001). The coupled oscillator model of between-hand coordination in alternate handtapping: A reappraisal. *Journal of Experimental Psychology: Human Perception & Performance*, 27, 251–265.