






Expressing Agreement in Swedish and Chinese: A Case Study of Communicative Feedback in First-Time Encounters

Anna Jia Gander^(✉) , Nataliya Berbyuk Lindström , and Pierre Gander 

University of Gothenburg, 405 30 Göteborg, Sweden
{anna.jia.gander,nataliya.berbyuk.lindstrom,
pierre.gander}@ait.gu.se

Abstract. This paper is a case study that explores how communicative feedback for expressing agreement is used by Swedish and Chinese speakers in first encounters. Eight video-recorded conversations, four in Swedish and four in Chinese, between eight university students were analyzed. The findings show that both Swedish and Chinese speakers express agreement more multimodally than unimodally. The Swedish participants most frequently use unimodal vocal-verbal feedback *ja* and *nej* (equivalent to *yes* and *no* in English respectively), unimodal gestural feedback, primarily head nod(s), multimodal feedback *ja* in combination with head nod(s) and up-nod(s). The Chinese participants most commonly use unimodal vocal-verbal feedback *dui* and *shi* (equivalent to *yes* in English), followed by the unimodal gestural feedback head nod(s), and multimodal feedback *dui* and *shi* in combination with head nod(s) to express agreement. Also, the findings indicate that the expression of agreement varies between both cultures and genders. Swedes express agreement more than Chinese. Females express agreement more than males. The results can be used for developing technology of autonomous speech and gesture recognition of agreement communication.

Keywords: Cross-cultural communication · Agreement · Communicative feedback · Swedish · Chinese · Cultural/gender differences · First-time encounters

1 Introduction

In social communication activities, understanding the counterpart's expressions of agreement through social signals (such as vocal-verbal unimodal feedback *hm*, *yeah*, *no*, gestural unimodal feedback a head nod/shake, or multimodal feedback *yes* in combination with a head nod) is a fundamental aspect for achieving mutual understanding in interaction. Expressing agreement indicates “the situation in which people have the same opinion, or in which they approve or accept something” (Cambridge University Press 2020). People can share convergent or divergent opinions, proposals, goals, reciprocal relations, attitudes and feelings through social signals of agreement and disagreement (Khaki et al. 2016). Recognizing and interpreting such social signals is an essential

capacity for social life, and for developing Intelligent Systems, from dialogue systems to Embodied Agents capable of sensing, interpreting and delivering social signals (Poggi et al. 2011). Thus, research on multimodal human interaction in naturalistic settings from various communication contexts is essential to enhance understanding of social signals and its applications in digital technology development.

Earlier research has shown that people express agreement in different ways, and culture is one of the factors which influences it. In high-context cultures such as China and Malaysia, face saving leads to avoidance of expressing disagreement and saying ‘yes’ doesn’t necessarily mean agreement, which is the opposite to low context cultures such as Germany and the U.S. (Hall 1959; Kevin 2004). The research of Swedes’ communication has shown that they give much feedback in interactions, signaling seeking consensus and avoiding conflict (Pedersen 2010). Gender is another factor influencing agreement expression. The research of Malaysian Chinese agreement strategies shows that females tend to express more agreement in interactions than the males (Azlina Abdul 2017).

Few studies in cross-cultural communication research focus on analysis of real-time interactions. Currently, there is little research with a focus on multimodal communication and cross-cultural comparison between Sweden and China, although they have increasing contacts and collaborations in both academia and industry, which motivates why more research is needed in this area.

2 Background

Social signals are perceivable stimuli that, either directly or indirectly, convey information concerning social actions, social interaction, attitudes, social emotions and social relations (Vinciarelli et al. 2012). Expressing agreement indicates “the situation in which people have the same opinion, or in which they approve or accept something” (Cambridge University Press 2020). Agreement and disagreement are expressed by specific social signals, through which people in the interaction express, for example, whether they share the same opinions, whether they accept each other’s proposals, whether they have convergent or divergent goals, attitudes and feelings, how their reciprocal relations are, and how to predict the development and outcome of the discussion or negotiation. When people argue in a discussion or negotiation, it is important for people to understand whether, what, and how they agree or disagree.

In order to perceive and understand the social signals of agreement and disagreement, one must be able to catch and interpret the relevant feedback expressions, which are often expressed in different communicative modalities – words, gestures, intonation, facial expressions, gaze, head movements, and posture (Gander 2018; Poggi et al. 2011). One may look at the other while smiling and nodding to express agreement, while shaking head and saying “no” to signal disagreement. Gestural unimodal expressions of agreement (such as head nod, finger thumb up) have been researched in different datasets such as TV live political debate (Poggi et al. 2011), dyadic dialogues about the likes and dislikes on common topics such as sports, movies, and music (Khaki et al. 2016), and human-computer interaction through gestures via touchless interfaces (Madapana et al. 2020). Expressions of agreement have also been explored for developing an implementation and training model of agreement negotiation dialogue, for example, with a

dataset of dialogues where people have contradictory communicative goals and try to reach agreement, in particular focusing on the sequence of communicative acts and the vocal-verbal unimodal means (Koit 2016). Mehu and van der Maaten (2014) stressed the importance of using multimodal and dynamic features to investigate the cooperative role of pitch, vocal intensity and speed of speech to verbal utterance of agreement and disagreement. Although researching both visual and auditory social signals of agreement and disagreement has obtained growing attention, studies on multimodal communication of agreement and disagreement still focus more on the role of nonverbal behavior in social perception (e.g., Mehu and van der Maaten 2014) rather than the combination of multimodalities in expressing agreement and disagreement. This paper will investigate both gestural unimodal, vocal-verbal unimodal, and their combined multimodal communicative behaviors to contribute to big data corpus of naturalistic interactions. Such empirical research findings can be used for developing and evaluating statistical models of agreement and disagreement in interactions for instance, and also for designing technology of autonomous speech and gesture recognition of agreement communication.

Besides modality issues, there have been also cultural and gender concerns of agreement communication. As known, understanding the counterpart's expressions of agreement is a fundamental aspect for achieving mutual understanding in cross-cultural communication activities. Across cultures, people express agreement in different ways. The research about the Malaysian Chinese agreement strategies shows that interlocutors tend to complete and repeat part of the previous speaker's utterance and give positive feedback, which is explained by striving for supporting speaker's face and harmony rooted in Chinese cultural values (Azlina Abdul 2017). The research also emphasizes that female speakers tend to express more agreement than male speakers. In high-context cultures, for example, China and Malaysia, face saving leads to avoiding expressing open disagreement with interlocutors and saying 'yes' doesn't necessarily mean agreement; whereas, a more direct communication like "no, I don't think so" or "sorry, I don't agree" is common in low context cultures, such as Germany and the U.S. (Hall 1959; Kevin 2004). The Swedes are known for giving much feedback in interactions, signaling consensus seeking and conflict avoidance (Pedersen 2010). Few studies in cross-cultural communication research focus on analysis of real-time interactions. Also, there is little research focus on multimodal communication and cross-cultural comparison between Swedish and Chinese communication patterns, though contacts between Chinese and Swedes are increasing. In recent years, Chinese direct investments in Sweden have tripled and are expected to increase further (Weissmann and Rappe 2017). Sweden and China have many international collaborations in both academia and industry. However, few studies of real-time interactions have assessed similarities and differences between Swedish and Chinese cultures in general, with little focus on expressing agreement in particular. We are going to address this in the present paper.

3 Research Aim and Questions

This paper aims at exploring and comparing how Chinese and Swedes express agreement in interactions. More specifically, it focuses on the multimodal communicative feedback for expressing agreement, including vocal, gestural and combination of vocal and verbal feedback expressions.

Two specific research questions will be addressed:

1. What feedback expressions are used for expressing agreement in the Swedish and the Chinese conversations?
2. What similarities and differences can be observed between the Swedish and the Chinese expressions of agreement?

4 Methodology

4.1 Data Collection and Participants

Our data consist of four Swedish-Swedish and four Chinese-Chinese first-time encounters' dyadic conversations between eight university students in Sweden. The conversations were in languages of Swedish and Mandarin Chinese respectively and were video- and audio-recorded in a university setting. Each conversation lasts approximately ten minutes. The participants were 23–30 years old, four males and four females. They were native speakers of Swedish and Mandarin Chinese. The participants did not know each other at the time of the recordings. They were invited to get acquainted and interact with one another, with no limitation on the topics. Anonymity was emphasized in the research project. The participants were asked for permission to use the pictures from the recordings for publications. All participants gave written consent for their participation in the research and the use of their data.



Fig. 1. Recording of the encounter.

Two participants (see, e.g., Fig. 1) were paired in a classroom and asked to interact with each other for about 10 min after the research assistant switched video-cameras on and left the room. The interactions were audio- and video-recorded by three video cameras (left-, center-, and right-position) with interlocutors standing face-to-face to capture the participant's communicative movement as much as possible. The cameras were placed off the participants' field of vision to minimize possible effects of video recording on participants' behavior.

Each participant was involved in two mono-cultural dialogues, in which the gender of the other interlocutor varied. The first seven minutes of each dialogue were studied, because the shortest dialogue was around seven minutes long. Within the time given, the use of agreement expressions was compared between the Chinese and the Swedish speakers. An overview of the data is presented in Table 1.

Table 1. Overview of the recordings (C = Chinese, S = Swedish, f = female, and m = male).

Recording	Participants	Analyzed length	Language
Dial.1	Cf1–Cf2	7:00 min. each (56 min in total)	Chinese
Dial.2	Cf1–Cm2		
Dial.3	Cm1–Cf2		
Dial.4	Cm1–Cm2		
Dial.5	Sf1–Sf2		Swedish
Dial.6	Sf1–Sm2		
Dial.7	Sm1–Sm2		
Dial.8	Sm1–Sf2		

4.2 Transcription, Coding, and Inter-coder Rating

The recordings were transcribed and annotated by native speakers of Swedish and Chinese, according to the Gothenburg Transcription Standard (GTS) version 6.4 (Nivre et al. 2004) and the MUMIN Coding Scheme for the Annotation of Feedback, Turn management and Sequencing Phenomena (Allwood et al. 2007). This allows inclusion of such features of spoken language as pauses and overlaps as well as comments on communicative body movements and other events. Feedback expressions through various modalities, their communicative act units, and their communicative functions were annotated. The following conventions from GTS and MUMIN were used in the paper (see Table 2).

Given the differences and connections between the perception modalities (e.g., visual modality, auditory modality, and haptic modality) and production modalities (e.g., gestural input is perceived through visual modality, vocal-verbal input through auditory modality, tactile input through haptic modality) (cf. Gander 2018), this paper focuses on the perception-sensory perspective. Thus, the researched agreement expressions are classified into three categories:

1. Vocal-verbal unimodal: verbal expressions only
2. Gestural unimodal: bodily communication, including head nods, facial displays, etc.
3. Vocal-verbal + gestural multimodal: combinations of (1) and (2)

The annotation was carried out by one annotator, and inter-coder reliability checking was performed independently by another annotator. The inter-coder agreement

Table 2. Transcription and MUMIN coding conventions.

Symbol	Explanation
\$Cf1	Speaker (Chinese, female)
[]	overlap brackets; numbers used to indicate the overlapped parts
{j}a	j is not pronounced
/, //, ///	a short, intermediate and a long pause respectively
:	lengthening
< >, @ < >	comments about non-verbal behavior, comment on standard orthography, other actions, clarifications
FB	feedback
VFB	vocal-verbal feedback
GFB	gestural feedback
CPUE/A	contact, perception, understanding, emotion/attitude
CPU	contact, perception, understanding

on the annotation of feedback and its communicative functions is moderate, with a free-marginal $\kappa = 0.58$.

By using the coding scheme (Allwood et al. 2007), the basic communicative function of feedback was coded as CPU (i.e., contact, perception and understanding). A code of CPU means ‘I want to and am able to continue the conversation and I perceive and understand what you have communicated.’ In addition, we also coded other communicative functions of feedback in our data, such as effective and epistemic stances which include emotion, attitude, agreement, acceptance, and evaluation etc. (Allwood et al. 2000; Paggio et al. 2017; Wessel-Tolvig and Paggio 2013; Gander 2018). There, agreement differs from acceptance in that agreement indicates acceptance, whereas acceptance does not indicate agreement. One can accept but may not necessarily agree. Acceptance is the necessary but not sufficient condition for agreement. In this paper, we focus specifically on agreement. CPU + agreement refers to that ‘I want to and am able to continue the conversation and I perceive and understand what you have communicated’ which is meant by CPU (only) plus ‘I agree with you on what you have said’.

During the coding work, we identified CPU + confirmation too, and distinguished it from CPU + agreement, with a consideration of whether it is confirming a truth or fact or agreeing on an opinion. We do not want to simplify a truth or a fact as an opinion, and we do not want to simply take ‘confirmation’ as one type of ‘agreement’. Accordingly, we define the criterion of identifying CPU + agreement and CPU + confirmation as below.

The code of CPU + agreement occurs when the interlocutor is agreeing on some subjective opinions that the other interlocutor has expressed in the previous utterance which is not necessarily the truth or fact; while, CPU + confirmation is coded when the interlocutor is confirming what the other interlocutor has said is true. As has been stated and reasoned above, in this research we do not consider confirmation as one type

of agreement. Instead, we take agreement to be based on shared views or opinions (not necessarily truth or fact), whereas confirmation to be associated with facts or truths.

4.3 Data Analysis

Occurrences of agreement feedback were identified and categorized according to modality (see Sect. 4.2, vocal-verbal unimodal, gestural unimodal, and vocal-verbal + gestural multimodal). Further, interactional sequences were qualitatively analyzed in order to examine the communicative functions of agreement feedback in the interactions. Additionally, quantitative analysis of frequencies and percentages of feedback occurrences were performed, comparing cultures (Swedish/Chinese) and genders. Because of the limited size of the samples, analysis of differences was made directly on the observed frequencies in the samples rather than being statistically tested, and generalizations were not attempted.

5 Results

The first research question aims to identify the agreement expressions in Swedish and Chinese. The second research question aims to explore the similarities and differences between the Swedish and the Chinese agreement expressions. Results are presented using raw frequencies and percentages of agreement feedback. Excerpts are extracted from the transcriptions.

5.1 General Overview of Agreement Feedback

This section presents a general overview of the types of agreement feedback in the data. Table 3 shows that the Swedish participants in the sample use more of both unimodal and multimodal agreement feedback than the Chinese participants (110 compared to 54 occurrences). Compared to the Chinese, the Swedes also use more vocal-verbal expressions of agreement (52 versus 17) and more multimodal feedback (41 compared to 27).

Regarding gender, both the Chinese and the Swedish female speakers use more agreement expressions than male (81% versus 19% in the Chinese sample and 67% versus 33% in the Swedish sample), including all the three types of modalities of agreement expressions. In the Chinese sample, the share of gestural unimodal and multimodal feedback expressions is much higher in the females than in the males (90% and 89% compared to 10% and 11% respectively). In addition, Table 3 also shows that the Swedish females use more agreement expressions than the Chinese females (74 compared to 44), so as the Swedish males compared to the Chinese males (36 versus 10).

Table 3. Overview of agreement feedback used by the Swedish and Chinese participants (FB = Feedback, M = Male, F = Female).

FB	Chinese			Swedish		
	M	F	Total	M	F	Total
Vocal-verbal unimodal	6 35%	11 65%	17 100%	15 29%	37 71%	52 100%
Gestural unimodal	1 10%	9 90%	10 100%	6 35%	11 65%	17 100%
Vocal-verbal + gestural multimodal	3 11%	24 89%	27 100%	15 37%	26 63%	41 100%
Total:	10 19%	44 81%	54 100%	36 33%	74 67%	110 100%

5.2 Swedish and Chinese Unimodal Vocal-Verbal Feedback Expressions for Agreement

Regarding unimodal agreement feedback expressions, we will look at the Swedish and the Chinese data respectively.

In the Swedish data, the identified Swedish unimodal vocal-verbal feedback expressions can be grouped into three categories: 1) *yes* and its variations, for example, *ja* (*yes*), *{j}a* (*yeah*), repetitions of *{j}a* (*yeah*); *{j}a: precis* (*yes exactly*); *jo* (*sure*); 2) *m* and its variations, for example, *m, m:* (lengthening); 3) *no* and its variations (e.g., when interlocutor agrees with preceding negative statement), such as *no, no no* (see Table 4). The Swedish participants show a tendency of expressing agreement using primarily *{j}a* and its variations, with an occurrence of 69% of all the unimodal vocal-verbal feedback expressions. The Swedish participants also use unimodal words *no* and *m* with their variations to express agreement, with frequencies of 17% and 14% respectively (see Table 4).

The data presented in Table 4 also shows that the Swedish females use more vocal-verbal feedback expressions of *yes* and *m* and the variations than the Swedish males, with occurrences of 50% versus 19% and 12% versus 2%, respectively. In the meanwhile, the Swedish females and males show similar preferences in using *no* and its variations, with frequencies of 9% and 8%.

To present an example from our data, in Excerpt 1 below, Swedish female 2 expresses agreement while listening to Swedish female 1 about Swedish female 1's experiences with home care. Swedish female 2 agrees with Swedish female 1 on that some people one meet at the workplace people can be nice while others can be avoided, by using unimodal vocal-verbal feedback expression *{j}a*.

Table 4. Unimodal vocal-verbal feedback (VFB) agreement expressions used by the Swedish participants

Swedish Unimodal VFB	English translation	M	F	Total
Yes and its variations				
<i>{j}a; {j}a:; ja; hja (ingressive); {j}a ja; {j}a {j}a;</i> <i>{j}a ja ja ja ja ja ja ja ja;</i> <i>{j}a eller hu{r}; {j}a: precis;</i> <i>precis;</i> <i>{j}a precis precis;</i> <i>{j}a precis jo verkli{j}en; {j}a jo precis; ja ja men precis; {j}a precis nä;</i> <i>hja exakt jo; kjo; {j}a jo</i>	<i>yeah; yeah:; yes; ha;</i> <i>yeah yes; yeah yeah;</i> <i>yes yes yes ... yes yes;</i> <i>eah that's right; yeah: exactly;</i> <i>exactly;</i> <i>yeah exactly exactly;</i> <i>yeah exactly sure that's true; yeah sure exactly; yes but exactly; yeah exactly no;</i> <i>yeah exactly sure; yeah sure</i>	10 19%	26 50%	36 69%
No and its variations				
<i>nä nä; näe; nä men visst; nä nä de{j} e klart; näe precis; nä</i>	<i>no no; no;</i> <i>no exactly; no no of course; no exactly</i>	4 8%	5 9%	9 17%
M and its variations				
<i>m; m:;</i> <i>m nä; m jo {j}a</i>	<i>yes; yeah:;</i> <i>yeah no; yeah sure yeah</i>	1 2%	6 12%	7 14%
Total:		15 29%	37 71%	52 100%

Excerpt 1. (SwDial.5) With some you work great, with others you don't.

- 1 \$Sf1 de beror ju så mycke på va man få fö utbyte me personen liksom så att
it depends so much on what kind of connection you get with the person like so that
- 2 / e // det e ju lite e e ja jobba hemtjänsten innan / å där e de ju också lite
/ eh // it's a little eh eh I worked home care before also / and there it's also a little
- 3 de att man träffar [väldigt mycke människor] / å vissa funkar man jättebra me
like you meet [1 a lot] of people / and with some you work great with
- 4 / å de e jättetrevlit // å andra e lite så hä bara / måste ja gå dit ida
/ and it's very nice // and others are a little like / do I have to go there today
- 5 \$Sf2 < [{j}a] >
< [1 yeah] 1 >
@ < vocal-verbal feedback; CPU + agreement >

Then, turning to the Chinese data, Table 5 presents the unimodal vocal-verbal feedback expressions used by the Chinese participants when expressing agreement.

Table 5. Unimodal vocal-verbal feedback (VFB) agreement expressions made by the Chinese participants.

Chinese Unimodal VFB	English translation	M	F	Total
Yes and its variations				
<i>dui, dui na shi</i>	<i>yes it is/ right</i>	4 23%	10 59%	14 82%
M and its variations				
<i>en, a</i>	<i>yeah; ah</i>	1 6%	1 6%	2 12%
No and its variations				
<i>bu shi</i>	<i>no, (it is) not</i>	1 6%	0 0%	1 6%
Total:		6 35%	11 65%	17 100%

Dui and its variation *dui na shi* (equivalent of *yes* and *right* in English) are the most common (82%) agreement expressions in the Chinese data. The Chinese female speakers give more agreement feedback than the Chinese males (65% compared to 35%).

Excerpt 2 below illustrates the Chinese female speaker 1 agreeing with the Chinese female 2 on that the work in China is faster tempo and more task oriented compared to that in Sweden:

Excerpt 2. (ChDial.1) Work tempo and task orientated.

- 1 \$Cf2 bu hui xiang zhe tou er gan jue guo de xiang bi jiao chong shi
it was not like here (I) feel life is more occupied and (task) oriented
- 2 \$Cf1 < **dui** > you ya li ma
< **right** > *pressure and drive isn't it*
@ < **vocal-verbal feedback; CPU + agreement** >

Comparing the data of Table 4 and Table 5, one can observe that in both Swedish and Chinese conversations the vocal-verbal feedback *yes* and its variations are the most common unimodal agreement expressions. Also, in both Swedish and Chinese data, the female speakers give more feedback than the males. Besides, there is more variety in feedback expressions used in the Swedish sample compared to the Chinese data.

5.3 Swedish and Chinese Unimodal Gestural Feedback Expressions for Agreement

The Chinese and the Swedish unimodal gestural feedback expressions for agreement will be analyzed in this section. An overview of the data is presented in Table 6.

Table 6. Swedish and Chinese unimodal gestural feedback (GFB) agreement expressions

Swedish Unimodal GFB	M	F	Total	Chinese	M	F	Total
a nod; nods	2	8	10	a nod; nods	0	8	8
shake	1	0	1	smile	1	1	2
Total	3	8	11	Total	1	9	10

As can be seen from Table 6, both Swedish and Chinese speakers use unimodal gestural expressions relatively seldom, compared to how they use unimodal vocal-verbal expressions (presented in the previous section). Nod(s) are the most common unimodal feedback expression to communicate agreement in both Swedish and Chinese samples.

Excerpt 3 below illustrates the Swedish male speaker 2 agreeing with the Swedish male speaker 1 on that travelling is better to do after having completed education. A possible reason for just nodding (of Sm2) in this excerpt can be that the speaker (Sm2) does not want to interrupt an apparently lively monologue of the other speaker (Sm1).

Excerpt 3. (Sw Dial.8) Travelling funds.

- 1 \$Sm1 nä de tycke inte ja bättre å göra nå du e klar kanske då
no I don't think so it's better to do when you're finished maybe then
- 2 å sen när du har gjort lumpen också får du lite pengar
and then when you have (...) done boot camp too you get a little money
- 3 där ifrån också så de har du mer i reskassa
from there too so it's eh do you have more travelling funds
- 4 \$Sm2 @ < head movement: nods >
@ < gestural feedback; CPU + agreement >

Similarly, in the Chinese data, by nodding the female speaker 2 shows agreement on that if there is an opportunity to work then one should work rather than studying (see Excerpt 4).

Excerpt 4. (Ch Dial.3) Work or study?

- 1 \$Cm1 < you ji hui jiu gong zuo >
< (if) there is an opportunity (I'll) work >
@ < gaze down to the left >
- 2 \$Cf2 < head nods >
@ < gestural feedback; CPU + agreement >

5.4 Swedish and Chinese Multimodal Feedback Expressions for Agreement

In this section, the Swedish and the Chinese multimodal feedback expressions for agreement will be presented respectively. We will start with the Swedish data, followed by the Chinese data.

The Swedish multimodal feedback expressions are grouped in the same categories as the vocal-verbal unimodal expressions (see Sect. 5.2), namely, (1) *yes* and its variations, (2) *no* and its variations, and (3) *m* and its variations. Bodily movements used with each vocal-verbal expression comprise the corresponding multimodal unit. Table 7 presents an overview of the Swedish multimodal expressions identified in the data.

Table 7. Multimodal feedback (VFB + GFB) agreement expressions used by the Swedish speakers.

VFB	English translation	GFB	M	F	Total
Yes and its variations + bodily movement					
<i>{j}a; ja; ja ja ja ja; ja det ä det; ja det är ja; jo</i>	<i>yeah; yes; yes yes yes yes; yes yes sure; it eh yes; sure</i>	nods	5	15	20
<i>ja; {j}a</i>	<i>yes; yeah</i>	up-nod	0	2	2
<i>{j}a</i>	<i>yeah</i>	tilts	4	0	4
<i>{j}a precis ja; {j}a {j}a precis absolut</i>	<i>yeah exactly yes; yeah yeah exactly absolutely</i>	nods	1	2	3
Subtotal:			10 24%	19 46%	29 70%
No and its variations + bodily movement					
<i>nä men man {j}a</i>	<i>no but you yeah</i>	nods	1	0	1
<i>näe näe näe exakt</i>	<i>no no no exactly</i>	tilt	1	0	1
<i>nä nä nä; m nä</i>	<i>no no no; yeah no</i>	shakes	0	2	2
<i>nä</i>	<i>no</i>	nod + tilt	0	1	1
<i>{j}a nä</i>	<i>yes no</i>	tilt + eyebrow rise	1	0	1
<i>nä de tycke inte ja</i>	<i>no I don't think so</i>	eyebrow frown	2	0	2
Subtotal:			5 12%	3 8%	8 20%
M and its variations + bodily movement					
<i>m; m ja; m jo; m yeah</i>	<i>yeah; yeah yes; yeah sure; yeah yeah</i>	nods	1	3	4
Subtotal:			1 2%	3 8%	4 10%
Total:			16 38%	25 62%	41 100%

Table 8. Multimodal feedback (VFB + GFB) agreement expressions used by the Chinese speakers

VFB	English translation	GFB	M	F	Total
Yes and its variations + bodily movement					
<i>dui; dui dui dui dui</i>	<i>right; right... right</i>	nods	0	2	2
<i>dui</i>	<i>right</i>	chuckle + nod	0	1	1
<i>dui</i>	<i>right</i>	smile + up-nod	0	1	1
<i>dui dui; dui</i>	<i>right right; right</i>	smile + nods	0	1	1
<i>a dui</i>	<i>ah right</i>	smile	1	0	1
<i>shi</i>	<i>yes</i>	nods	0	1	1
<i>en shi</i>	<i>yes it is</i>	nods	0	1	1
<i>ying gai shi</i>	<i>should be</i>	nods	0	1	1
<i>en</i>	<i>yes</i>	nod(s)	0	11	11
<i>en</i>	<i>yes</i>	chuckle + nods	0	1	1
Subtotal:			1 5%	20 90%	21 95%
No and its variations + bodily movement					
<i>bu shi</i>	<i>no/not</i>	shake	0	1	1
Subtotal:			0 0%	1 5%	1 5%
Total:			1 5%	21 95%	22 100%

Excerpt 6 presents an example of the Chinese female speaker 2 agreeing with the Chinese male speaker 1 on the open mindedness of their families regarding what they are allowed to do or not to do. The speaker Cf2 first provided a verbal feedback *en* (equivalent to *yeah* in English), followed by head nods and a lengthy pause which reinforced her vocal-verbal expression of agreement.

Excerpt 6. (Ch Dial. 3) Openness of families.

- 1 \$Cm1 dan shi wo shi mei you wo men jia dou hen kai fang
but I do not have my families are open-minded
- 2 \$Cf2 < en /// > wo ma ye shi peng you ///
 < yes /// > my mother has a friend (like this) too ///
 @ < head nods >
 @ vocal-verbal and gestural feedback; CPU + agreement

6 Discussion

Regarding modalities, the data has shown that both the Swedish and the Chinese speakers use vocal-verbal unimodal, gestural unimodal and multimodal expressions for expressing agreement. Unimodal gestural feedback expressions (e.g., head, shake, and smile) are less frequently used, compared to unimodal vocal-verbal and multimodal feedback expressions. This possibly indicates that vocal-verbal means are prioritized over the gestural means when people express agreement, and that simply nodding, for example, is not enough to ensure or reinforce agreeing in both cultures. Nevertheless, in both Swedish and Chinese samples, nodding is still the most common unimodal gestural expression of agreement, which is in line with previous research (Andonova and Taylor 2012). The findings also show that both Swedish and Chinese participants use head nod(s) in combination with *yes* of their own respective languages as the most common multimodal feedback expressions for showing agreement, which is consistent with earlier research (Lu and Allwood 2011). Also, the Swedish participants use more agreement feedback expressions both unimodally and multimodally than the Chinese participants. This is consistent with the prior studies which have noted the importance of consensus in Sweden (Havaleschka 2002) and the tendency that giving much feedback in interactions indicates listening and agreeing (Allwood et al. 2000; Tronnier and Allwood 2004) and avoiding conflict (Daun 2005; Pedersen 2010).

Besides the features of modalities for expressing agreement, both Swedish and Chinese female participants in this research use more agreement feedback than the males. This finding supports the research of Deng et al. (2016) on that females tend to have higher emotional expressivity and provide more feedback than males. In addition, although it has been found that both the Swedish and the Chinese female speakers express agreement more often than the males, this gender difference was bigger in the Chinese sample compared to the Swedish sample. This finding is in line with that of Azlina Abdul's (2017) research, in which the Chinese females are found giving considerably more agreement feedback than the Chinese males. Further, as Sweden is considered to be a society in which gender roles overlap to a higher extent than in China, it might be reflected in more similarities between the Swedish male and female participants in expressing agreement (Hofstede 1994).

The data has also shown that the Chinese participants provide considerably fewer agreement feedback expressions than Swedes. It can potentially be explained by that in Sweden interactions are often informal in particular among students at a university context, due to a relatively low power distance (Hofstede 1994). On the contrary, in China, a differentiation is clearly made between polite/impolite and insider/outsider communication, the former often occurring between strangers while the latter between close acquaintances (Fang and Faure 2011). As Chinese culture is a high context culture, many Chinese people are cautious in communicating with 'outsiders' by, for example, limiting their contact to brief and functional communicative exchanges (Hall 1959), which might be reflected in the Chinese speakers in our research being less willing to express themselves compared to the Swedes.

Because these findings are based on a particular communication context of first encounters and between a limited number of participants, the present research can be seen as a case study for the future work with more participants and varied contexts to

investigate agreement expressions, patterns and strategies in Swedish and Chinese or Western and Eastern cultures further.

7 Conclusion

This paper explored and compared how Chinese and Swedes express agreement in first-time encounter's conversations by using social signals of communicative feedback. In this paper, eight audio- and video-recorded interactions between four Swedish and four Chinese participants were analyzed. The identified agreement feedback expressions were classified into three types according to which modalities are involved, namely, the vocal-verbal unimodal, the gestural unimodal, and the vocal-verbal + gestural multimodal expressions. Two research questions were investigated, regarding respectively what feedback expressions are used for expressing agreement in the Swedish and the Chinese conversations and what similarities and differences can be observed between the Swedish and the Chinese expressions of agreement.

About the feedback expressions used for expressing agreement, we have found that participants from both samples use both unimodal and multimodal feedback expressions. The Swedish participants most frequently use unimodal vocal-verbal feedback *ja* and *nej* (equivalent to *yes* and *no* in English respectively), unimodal gestural feedback, primarily head nod(s), multimodal feedback *ja* in combination with head nod(s) and up-nod(s). In the Chinese data, the most commonly used feedback is unimodal vocal-verbal expression *dui* and *shi* (equivalent to *yes* in English), followed by the unimodal gestural expression head nod(s), and multimodal expression *dui* and *shi* in combination with head nod(s) to express agreement.

Regarding the similarities between the Swedish and the Chinese samples, this case study suggests that both Swedish and Chinese speakers express agreement more multimodally than unimodally, most often with *yes* in combination with nod(s). The infrequent use of unimodal agreement expressions possibly indicates that simply nodding, for example, is not enough to reinforce agreeing. Also, in both samples the females showed more agreement than males. On the differences between the Swedish and the Chinese data, the Swedish participants expressed more agreement than the Chinese, which can be explained by the value the Swedes put on reaching consensus and avoiding conflict. Also, the gender difference between males and females in expressing agreement has been found bigger in the Chinese sample than the Swedish one. This might be because that Sweden is considered to be a society in which gender roles overlap to a higher extent, which could be reflected in more similarities between male and female swedes in social behaviors.

This case study suggests that the expression of agreement varies between cultures and genders. Both Swedish and Chinese speakers express agreement more multimodally than unimodally. Females express agreement more than males. Swedes express agreement more than Chinese. In spite of the limitations of single communication context and number of participants, this paper contributes to an increased understanding of Swedish and Chinese cultural similarities and differences in expressing agreement. The empirical findings can contribute to cross-cultural multimodal communication research and intercultural practices such as business negotiation and educational collaboration. It can

also contribute to big data corpus and automatic analysis of agreement communication, which may help designing technology of autonomous speech and gesture recognition, and developing machine socio-cultural training.

Acknowledgements. The authors would like to thank the participants for their time and engagement.

References

- Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., Paggio, P.: The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Lang. Resour. Eval.* **41**, 273–287 (2007)
- Allwood, J., Traum, D., Jokinen, K.: Cooperation, dialogue and ethics. *Int. J. Hum.-Comput. Stud.* **53**, 871–914 (2000)
- Andonova, E., Taylor, H.: Nodding in dis/agreement: a tale of two cultures. *Cogn. Process.* **13**, 79–82 (2012)
- Azlina Abdul, A.: Agreement strategies among Malaysian Chinese speakers of English. *3L, Language, Linguistics, Literature*, 23 (2017)
- Cambridge University Press: Cambridge Dictionary Online (2020)
- Daun, Å.: En stuga på sjätte våningen: Svensk mentalitet i en mångkulturell värld. B. Östlings bokförlag Symposion, Eslöv, Sweden (2005)
- Deng, Y., Chang, L., Yang, M., Huo, M., Zhou, R.: Gender differences in emotional response: Inconsistency between experience and expressivity. *PLOS ONE* **11**, e0158666 (2016)
- Fang, T., Faure, G.O.: Chinese communication characteristics: a Yin Yang perspective. *Int. J. Intercult. Relat.* **35**, 320–333 (2011)
- Gander, A.J.: Understanding in real-time communication: micro-feedback and meaning repair in face-to-face and video-mediated intercultural interactions, Doctoral dissertation, University of Gothenburg (2018)
- Hall, E.T.: *The Silent Language*. Doubleday, Garden City (1959)
- Havaleschka, F.: Differences between Danish and Swedish management. *Leadersh. Organiz. Dev. J.* **23**, 323 (2002)
- Hofstede, G.: Cultures and organizations: software of the mind. *Organiz. Stud.* **15**(3), 457–460 (1994)
- Kevin, A.: Culture as context, culture as communication: considerations for humanitarian negotiators. *Harv. Negot. Law Rev.* **9**, 391–409 (2004)
- Khaki, H., Bozkurt, E., Erzin, E.: Agreement and disagreement classification of dyadic interactions using vocal and gestural cues. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, pp. 2762–2766 (2016)
- Koite, M.: Developing a model of agreement negotiation dialogue. In: *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2016)*, pp. 157–162. SCITEPRESS – Science and Technology Publications, Setubal, Portugal (2016)
- Lu, J., Allwood, J.: Unimodal and multimodal feedback in Chinese and Swedish monocultural and intercultural interactions (a case study). *NEALT Proc. Ser.* **15**, 40–47 (2011)
- Madapana, N., Gonzalez, G., Zhang, L., Rodgers, R., Wachs, J.: Agreement study using gesture description analysis. *IEEE Trans. Hum.-Mach. Syst.* **50**(5), 434–443 (2020)
- Mehu, M., van der Maaten, L.: Multimodal integration of dynamic audio-visual cues in the communication of agreement and disagreement. *J. Nonverbal Behav.* **38**(4), 569–597 (2014)

- Nivre, J., et al.: Göteborg transcription standard. V. 6.4. Gothenburg: Department of Linguistics, Göteborg University (2004)
- Paggio, P., Navarretta, C., Jongejan, B.: Automatic identification of head movements in video-recorded conversations: can words help? In: Proceedings of the Sixth Workshop on Vision and Language, pp. 40–42. Association for Computational Linguistics, Valencia, Spain (2017)
- Pedersen, J.: The different Swedish tack: an ethnopragmatic investigation of Swedish thanking and related concepts. *J. Pragmat.* **42**, 1258–1265 (2010)
- Poggi, I., D’Errico, F., Vincze, L.: Agreement and its multimodal communication in debates: a qualitative analysis. *Cogn. Comput.* **3**, 466–479 (2011)
- Tronnier, M., Allwood, J.: Fundamental frequency in feedback words in Swedish. In: The 18th International Congress on Acoustics (ICA 2004), pp. 2239–2242. International Congress on Acoustics, Kyoto, Japan (2004)
- Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D’Errico, F.: Bridging the gap between social animal and unsocial machine: a survey of social signal processing. *IEEE Trans. Affect. Comput.* **3**(1), 69–87 (2012)
- Weissmann, M., Rappe, E.: Sweden’s approach to China’s belt and road initiative: still a glass half-empty. The Swedish Institute, Stockholm, Sweden (2017)
- Wessel-Tolvig, B.N., Paggio, P.: Attitudinal emotions and head movements in Danish first acquaintance conversations. In: 4th Nordic Symposium on Multimodal Communication, pp. 91–97. Northern European Association for Language and Technology, Gothenburg, Sweden (2013)