Effects of Computer-Mediated Communication on Backchannel Responses : A Comparison across Cantonese, English, Japanese

Wei Yeng Shek, Wendy
The Hang Seng University of Hong Kong

1 Introduction

Communication is vital in facilitating the exchange of information and the reduction of intimacy between people. With the advent of the internet, more people substitute traditional face-to-face (FTF) communication with computer mediated mediums like Zoom and Skype. The significance of this study could be understood in two ways. Even backchannels are historically neglected in past lexical and discourse research, they are among the most effective measurements of the communicative efficiency in a speech community, indicating its research values on conversational interactions. Moreover, even though backchannels in different languages share the function of creating a bonding effect among speakers (Kraaz & Bernaisch, 2020), the behavior is likely to vary by the context as well as the language community. For example, it is reported that speakers of American English have higher use of backchannels in general than speakers of Mandarin Chinese (Tao & Thompson, 1991; Clancy et al., 1996). The increase in the use of CMC has presented a research gap that concerns not only the differences in cultures but also the differences in modalities. This research examines the usage of backchannels during Instagram lives across languages, allowing us to understand the cross-linguistic and cross-modal effects on conversational behaviors.

2 Literature Review

2.1 *The use of Computer Mediated Communication* Our communication forms are constantly refined by technology. Computer-Mediated Communication (CMC) is a broad term that refers to communication through interactive media. Described by Romiszowski and Mason (1996), it is the process by which people "create, exchange, and perceive information through network telecommunications". However, electronic media of communication often provide less communication stimuli to speakers, making interpretation of meaning more difficult. The Media Naturalness Theory developed by Kock (2005) argues that the human brain has evolved to develop a biological communication apparatus that processes a brain circuit that specializes in FTF interactions (Kock et al., 2007; Shi, 2020). Based on the features of FTF communication, one can evaluate the naturalness of a medium by the following five components:

(1) Co-location that allows speakers to see and hear each other
(2) Synchronicity, in which speakers could exchange communicative stimuli quickly
(3) Ability to observe facial expressions and react to it
(4) Ability to observe body languages
(5) Ability to convey and to listen to speeches

In other words, the higher the similarity the medium is to FTF interactions, the higher the naturalness of the conversation. Since visual-based CMCs can imitate the conversational features in FTF settings, it has become an increasingly popular substitute for FTF communication under the pandemic. Even though more non-verbal cues are conveyed in visual-based CMCs, there are still certain differences like frame sizes and network connection that affects the smoothness of the conversation (Passarelli, 2020). To assess the naturalness of computer-mediated conversations, we measure backchanneling behaviors.

*I would like to thank my advisor Mr. Charles Lam for the encouragement and support throughout the preparation for the Asian Junior Linguistic Conference. I would also want to show gratitude to all the AJL committee for their efforts to organize the conference.

2.2 *Definition of Backchannel* Members in a language community share sets of communication devices to enable synchronized communication. Listener responses, such as utterances like *yeah* and *okay*, are often used to show attentiveness towards the primary speaker. Yngve (1970) coined these listener responses by the term "backchannels", proposing the idea that the speaker and listener (also named non-primary speakers) engage in both speaking and listening roles during a conversation**.** White (1989) further elaborates Yngve's interpretation by suggesting that the speaker is in the primary channel (front channel), and the non-primary channel (back channel) consists of utterances generated by the listeners as responses. While evaluating backchannels, only the responses collected in the non-primary channel are considered. Both Yngve (1970) and White (1989) emphasize that backchannels do not take the floor while conveying useful linguistic information. Duncan and Fiske (1977) extended Yvgne's definition from a one-word level to include sentence completions and brief statements. Besides verbal responses, non-verbal responses like nodding and laughing are also taken into account when explaining backchannels (Mizutani, 1983).

While reviewing studies that examine responses of non-primary speakers, we have discovered different terminologies and interpretations of backchannels. Schegloff (1982) termed the phenomenon "continuer" by its function of moving the conversation forward. As functions of the backchannel change according to the actual contexts, it is difficult to select a particular term to represent all utterances. Another approach to defining backchannels is emphasizing the lexical length of the utterance**,** i.e., Minimal Token (Fujimoto, 2009). However, as backchannel size can be varied differently in different contexts, it is difficult to generate a standardized definition. Tottie (1991) considers expressions with more than one item like yeah yeah right as a single backchannel, while Clancy et al. (1996) identify the same utterance as three separate backchannels since its definition only includes non-lexical single word utterances. As the above definitions could only express part of the backchannel features, in this project, we use the term "backchannel" to refer to listener responses as it is widely used in similar studies. Backchannels will be classified through three significant criteria:

(1) Useful information is conveyed, in terms of keeping the conversation going (Yngve, 1970).
(2) There should not be any intention to claim the floor of the speaker (White, 1989; Tao and Thompson, 1991)
(3) The lexical length should be minimized (Peters & Wong, 2006).

We will also examine various kinds of listeners' responses, including non-lexical utterances and gestures as long they fulfill the criteria above.

2.3 *Functions and expectations across language*

2.3.1 *Functions* Based on interpretations by different scholars, Maynard (1991) and Schegloff (1982) summarized the functions of backchannels and categorized them into six types.

(1) Continuer: the listener encourages the primary speaker to continue even if they had the opportunity to initiate a full turn.
(2) Understanding: the listener feels the need to assure that they comprehend the primary speaker's information.
(3) Agreement: the listener reacts to previously known information made by the speaker.
(4) Support and Empathy: the listener responds by showing support or empathy towards an evaluative judgment.
(5) Emotive Responses: the listener answers emotionally to the primary speaker's statement.
(6) Minor Additions: the listener intends to correct or add something to complete the utterance of the primary speaker.

2.3.2 *Expectations on the Frequency of Japanese backchannels* In previous comparative studies, Japanese speakers were reported to produce more backchannels than their English and Mandarin speaking counterparts (Clancy et al., 1996; Cutrone, 2005). From the linguistic point of view, the brief pauses which are marked by final particles in Japanese language provide a convenient environment for its speakers to generate more backchannels (Maynard, 1997; White, 1989). Japanese speakers would use なるほど (Naru Hodo) and 確かに (tashikani) to express their understanding towards the speakers, while English and Cantonese speakers would use simpler and shorter utterances (e.g *yeah* and *hai6*) to illustrate the same meaning. From a cultural perspective, the Japanese language is known for its indirectness and politeness to establish a closer relationship with each other. They are expected to be considerate and be very sensitive to the listeners' emotions and feelings to maintain harmony and unanimity. Giving frequent responses to the primary speaker can show constant interest and empathy to the speaker, even though sometimes it involves the compliance of opposed ideas (White, 1989). Therefore we expect to receive more listener responses in Japanese, compared to the production of the other two languages.

2.4 *Classification of backchannels* Backchannels could be further categorized into different subterms (Clancy et al., 1996; Lempinen, 2020).
(1) Reactive Tokens are short non-lexical responses that serve the function of understanding and agreement.
(2) Reactive expressions exist in the form of phrases or short statements, usually emotive expressions.
(3) Collaborative finishes occur when the listeners react by finishing the speaker's utterances.
(4) Repetitions occur when the listener reacts to the speaker's statement by repeating part of the previous utterances.
(5) Resumptive Openers are structurally Reactive Tokens used at turn-initial points, which are usually followed by short pauses.

3 Research Questions

a) Are different backchannel frequencies and patterns of distributions displayed in CMC ?
Hypothesis 1: Speakers in CMC are predicted to produce a lower frequency of verbal backchannels and the lexical length of the produced backchannels will be longer.

b) How would turn-taking behaviours be affected by CMC media?
Hypothesis 2: Speakers in CMC media are expected to display less turn-taking strategies to allocate turns.

4 Methodology

4.1 *Data Collection* Three video-taped discourses from English, Cantonese, and Japanese native speakers are examined. The video clips[1] were all extracted from Instagram lives, which in total 6 males, with age range within the 30s were involved in the study. To eliminate the variations in communicative skills and formality, each conversation was carried out by 2 speakers in a non-argumentative, informal setting. The topics of their conversation are their quarantine lives. Since we focus on backchannel behaviors in conversations, only clips involving dialogues will be examined. Therefore, monologues and audience interactions will be deducted from the videos and excluded from our study.

5 Findings

5.1 *Frequency* Table 1.1 shows the frequency of backchannels in both modalities. In order to compare the changes in productions in different modes, we have also included references of verbal backchannels from previous research. As conversation research in Cantonese is scarce, this study uses Mandarin Chinese, which is similar to Cantonese in various aspects. In the FTF medium, the mean number of verbal responses is 5.3 per minute. In our data, the number of verbal backchannels in English and Cantonese CMC has increased, while our Japanese data has drastically decreased. However, the frequency of verbal backchannels under CMC are still highest in Japanese discourses, where a frequency of 8.7 responses per minute is displayed. In addition to verbal responses, we have also collected data on non-verbal backchannels. The usage of non-verbal backchannels like nodding and laughters are mostly displayed in Cantonese, with an occurrence of 6.3 responses per minute on average.

[1] The three clips are all approximately 20 minutes in length.
Cantonese (23m 57s): [https://www.youtube.com/watch?v=A8PGU5dPPP8&t=245s]
English (24m 36s): [https://www.youtube.com/watch?v=ZmP7ebgFSL0&t=667s]
Japanese (19m 32s): [https://www.youtube.com/watch?v=Vc5C80Xw0Rc&t=309s]

	English	Japanese	Cantonese
FTF Verbal Backchannels [2]	2.8/ min	13.33/ min	1.3/ min
CMC Verbal backchannels	5.6/ min	8.7/ min	1.62/ min
CMC Non Verbal backchannels	2.8/ min	1.8/ min	6.3/ min

Table 1.1 *Frequency of Backchannels Per Minute*

5.2 *Variation* There were several types of backchannels displayed in our col-lected data. In this section, to compare the exact distribution of verbal backchannel types, we provide an analysis on their usage with reference to the classification of backchannels by Clancy et al. (1996), and Lempinen (2020). Table 1.2 compares the distribution of verbal responses between FTF and CMC. As the exact data size of the research is not specified in their prior research, we will be converting our collected data in percentage to achieve a clearer illustration. The data shows that Reactive Tokens are the most frequent type of backchannels found across three languages. Even though there is a slight decrease in the Reactive Token percentage in Japanese (68.3% →56.7%), the percentage in English and Cantonese has increased nearly a double. In Cantonese, the percentage of Reactive tokens covers approximately 80% of the total occurrence. In contrast, the use of Reactive expressions in CMC has decreased significantly in English and Cantonese, which the drop is more significant in our Cantonese data, with a percentage of 7.8%, compared to 31.3% in the FTF medium. The occurrences of Repetitions and Resumptive openers has decreased in CMC across three languages.

Language	Modality	Reactive Tokens	Reactive expressions	Collaborative finishes	Repetitions	Resumptive Openers
Japanese	FTF	68.3%	17.0%	0%	2.2%	12.5%
	CMC (n=104)	56.7%	35.6%	0.96%	3.8%	2.9%
English	FTF	37.9%	34.2%	15.6%	1.3%	10.4%
	CMC (n=74)	71.6%	18.9%	1.35%	1.4%	8.1%
Cantonese	FTF (Mandarin)	47.2%	31.1%	8.9%	5.8%	14.5%
	CMC (n=51)	80.3%	7.8%	3.9%	2.0%	5.9%

Table 1.2 *Types of Backchannels in Japanese, English, and Cantonese*

Reactive Tokens (RT) : Non-lexical responses that display the listener's interest or understanding towards the primary speakers.

[2] The verbal backchannels frequency in FTF medium are based on prior research and the data are converted to per minute for better comparison. Japanese [Hanzawa,2012], English [White, 1989], Cantonese [Wan, 2018]

Common RTs in Japanese, English and Cantonese

Japanese: うん[unn], ええ[ee], はい[hai], へえ[he]
English：Yeah, Yep, Hmm, Oh
Cantonese: 係[hai2], 唔[m4], 哦[Ngo4]

RTs do not only exist in singular forms, in the examples below, CMC speakers display types of responses by duplicating frequent used RTs to form backchannels with a lexically longer length[3].

Cantonese

(1) A: 其實我係未出碟之前已經係鋪到明, 就話呢個 concept 就^叫做 negative.
kei4 sat6 ngo5 hai2 mei6 ceot1 dip2 zi1 cin4 ji5 ging1 hai6 pou1 dou3 ming4 , zau6 waa6 ni1 go3 concept zau6 giu3 zou6 negative.
'Actually I have already give clear hints before releasing the album, which we named the concept as "negative".'
B: ^係係係. [nodding]
hai6 hai6 hai6.
'Right right right' [nodding]

Japanese

(2) A: 僕もちょっど頭で少し喋りましたけど　その先日発表になったアラフェス？
boku mo choddo atama de sukoshi shaberimashitakedo sono senjitsu happyō ni natta arafesu ?
'Although I've mentioned a bit in the beginning, the Arafes (concert) was announced a few days ago?'
B: はいはいはい.
hai hai hai
Yeah yeah yeah.'

In example (3), instead of the duplication of RTs, CMC speakers in English has co-join REs to form a longer backchannel.

English

(3) A: The guys are actually teaching and I'm like..oh these guys have done that several times.
B: **That's right.**[Nodding]**That's right.**

Reactive Expressions (RE) : Non floor-taking responses mostly function as expressions of emotions, but sometimes they could also be continuers.

Common REs in Japanese, English and Cantonese

Japanese: そうだね[Soodane], そうですね[soodesune], なるほど[naruhodo], いいね[iine]
English: Oh yeah, Really, Oh my god, wow
Cantonese: 嘩係[waA hai2], 係[hai6], 唔[m4]

Examples (4) and (5) illustrate how backchannels could exist in question forms. Although the responses of B and B seem to be a question, the answers are not addressed in the following turn. Therefore, we categorize these utterances similar to responses like *Oh really* in English.

Cantonese

(4) A: 係我個同事 呀 Robert 去 A＆R 去俾呢隻歌我.^咁啱啱開始嘅時候...
hai6 ngo5 go2 tung4 si6 aA Robert heoi3 A&R heoi3 bei2 ni1 zek3 go1 ngo5 ^gam2 aam1 aam1

[3] These responses are counted as one Reactive Token (Instead of three) in our analysis as they serve the same purpose of showing attentiveness even if they do not share the same phonetic form.

hoi1 ci2 ge3 si4 hau6
'It was my colleague Robert who went to A&R and gave me this song..So when we first start..'
B: ^哦 都幾年啦?
o2 dou1 gei2 nin4 laa4 ?
'Oh that was years ago?'

Japanese
(5) A: でも 見ていましたよ^ 楽しそうだね
'Demo miteimashitayo tanoshi sōdane
'But I have watched it. It seemed interesting.'
B: ^あ、本当？
^a, hontō
'Oh, Really?

Collaborative Finishes : they occur when listeners finish the utterances of the speakers. To Japanese speakers, this specific type of listener utterances was called "*Sakidori*" (Horiguchi,1988; Im and Lee, 1995), which by literal translation, means taking advance, implying that non-primary speakers has to anticipate the speaker's words in order to produce relevant utterances. Collaborative finishes are previously reported to only occur in English and Mardarin, but not in Japanese. However, our study was able to gather examples from Japanese CMC speakers. Example (5) demonstrates the use of both Reactive tokens and Collaborative Finishes in his response. At first, B used はい to encourage A to continue his turn. However, when A struggles to address the object he was intended to refer to, B takes advance and finishes his sentence.

Japanese
(6) A：あのマイケル・ジャクソンのさ^ キラキラの ^手袋
ano maikeru jakuson no sa ^ kirakira no ^ tebukuro
'The shiny thing that Michael Jackson wears ^ gloves
B: ^はい[Nodding]^手袋ね...
^hai [Nodding]^ *tebukurone ..*
'Yeah [Nodding]...^The gloves right.'

Besides completing the utterance in a lexical level, we also observed CMC speakers also would predict the reactions of speakers as a form of collaborative finishes. In example (8), B is trying to imitate the possible reactions of A. Similarly, Cantonese CMC speakers complete A's utterance by predicting the reactions of the people who A referred to.

English
(7) A: And I'm like yeah, okay, We'll watch it...^
B: ^You are like, **wait, is that you? is that you**?

Cantonese
(8) A: 呢樣嘢出自我把口 佢哋兩個都會^
Ni1 joeng6 je5 ceot1 zi6 ngo5 baA hau2 keoi5 dei6 loeng5 go3 dou1 wui5
'When it comes out from my mouth, both of them are'
B: ^哦..咁嘅反應
^o4 ..gam3 ge3 faan2 jing3
'And they're like "Oh" '

Repetitions: coded when listeners repeat the utterances of the primary speaker. Example (9) shows how listeners respond by repeating a portion of the previous speech.

Cantonese

(9) A: 有啲係投咗降 有一兩個係 啊 真係搞唔掂.
jau5 di1 hai2 tau4 zo2 hong4 jau5 jat1 loeng5 go3 hai2 aA zan1 hai2 gaau2 m4 dim6
'Some of them gave up, one or two of them are like ..oh...that's really tough
B: **真係搞唔掂.**
zan1 hai2 gaau2 m4 dim6
'Yeah, it's really tough'

In example (10), A expressed uncertainty about the exact number of citizens, instead of just responding with a simple reactive token, B confirms A's statement by repeating the correct number. Therefore, we could see how repetitions could also consolidate the answers of the primary speakers.

English

(10) A: For those who don't know the selection, it takes either forty citizens?^ To the SEAL team training.
B: ^Forty for sure.

5.2.1 *Non Verbal Backchannels* Data of non-verbals backchannels are also examined in our study. From table 1.3, nodding is the most frequent backchannel displayed in the CMC medium across three languages, followed by laughter. Cantonese has displayed the highest frequency in both forms as well as the total number of non-verbal backchannels. Moreover, non-verbal responses can be combined with verbal backchannels. While Japanese speakers tend to combine nodding with Reactive Expressions, English and Cantonese speakers combine nodding with Reactive Tokens.

	Japanese	English	Cantonese
Nods Nods +____ Laughter Facial Expressions Gesturing Head Shake	16 1 (RE) 5 1	26 3 (RT) 8 2 3 9	82 5 (RT) 9 1
Total Number	23	51	97

Table 1.3 *Frequency distribution of Non-verbal backchannels*

English has displayed the most dispersed types of non-verbal backchannels in our collected data, including facial expressions and gesturing. In Fig. 1.1, the listener used gestures to give non-verbal feedback to the primary speaker. However, since the camera is not able to fully capture his hands in the video frame, there is an ambiguity in understanding the intended meaning of the gesture: It could be interpreted "yeah" , "exactly" or even "I told you so", which could either fit into this particular context. Therefore, from this example, we could see the use of CMCs cause difficulties in decoding non-verbal reactions, which eventually lead to misinterpretations of messages.

English

(11) A: I was trying to really convince them to sit early , coz like every other comment was like, "release it now release it now"

B:  (Gesture with hands)

Fig. 1.1 *Gesture Response of the English Speaker*

This graph illustrates the distribution of the two types of backchannels in the three languages. Japanese has the highest frequency of verbal backchannels and Cantonese has the highest frequency of non-verbal backchannels. Moreover, the use of non verbal and verbal backchannels are in a negative relationship, where the use of either one type compensate the other one. The data also display that CMC speakers would develop a usage pattern on the types of backchannel used during conversations, and the choice of these types are culturally specific.

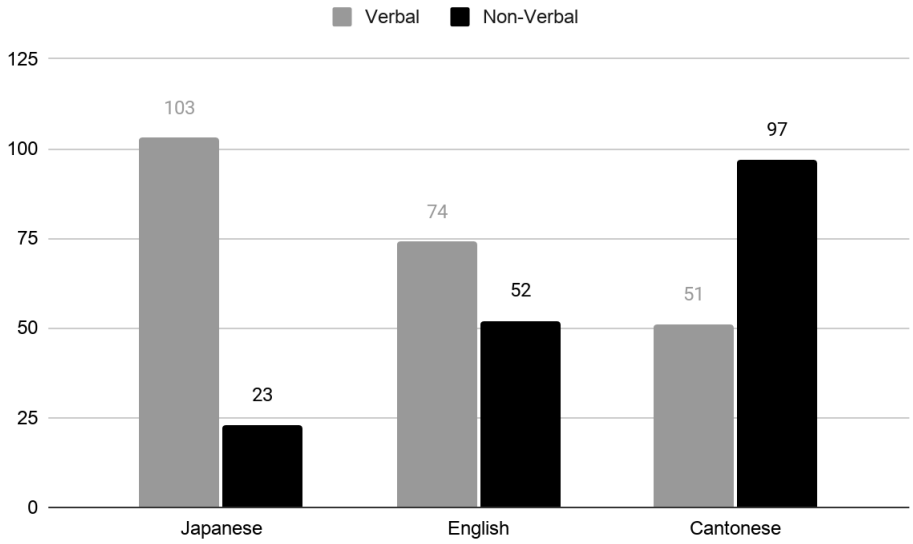

Fig. 1.2 *Comparative Graph of the Use of Verbal and Non-Verbal Backchannels*

5.3 *Turn-Taking* The smoothness of a conversation could be measured by the use of Resumptive openers.Resumptive openers are used at turn-initiated points to illustrate full turns. With a reduction in the percentage of resumptive openers (Table 1.1) as well as others turn-taking signals, it might be difficult for speakers to notice potential turns. In the conversations, we observed various situations of overlaps and silences.

5.3.1 *Overlaps* Overlapping is commonly observed during turn-taking behaviours. There are two types of overlappings found in our data: Cooperative overlaps and competitive overlaps function differently in terms of their natures. Co-operative overlaps a supportive nature that upholds the floor-right of the primary speaker. In example (12), B was giving additions to further illustrate A's description on the book's length.

Japanese

(12) **A:** そんなに^長くないからね
sonnani nagakunaikarane.
'Since it (The book) is not that long'
B: **^**ぎゅっと 集中する いけるのよ
^gyutto shūchū suru ikerunoyo
'If we really concentrate, we can finish it.'

Overlaps can exist in the form of repetitions. In this example, B was overlapping in order to encourage A to further develop on this topic.

Cantonese

(13) A: 之前青春頌嘅許廷鏗^去咗邊度?
zi1 cin4 cing1 ceon1 zung6 ge3 heoi2 ting4 hang1^ heoi3 zo2 bin1 dou6?
'Where is the Alfred who sang "Ode to Youth"'
B: ^好正面嗰個去咗邊度呢?
^Hou2 zing3 min6 go2 go3 heoi3 zo2 bin1 dou6 ne1?
'Where's that positive Alfred?'

Therefore, as cooperative overlaps do not claim the floor, they are seldom perceived as intrusive, thus speakers in CMC tend to tolerate such overlaps in utterances and continue their conversations.

Competitive overlaps have a disruptive nature. Instead of focusing on the primary speaker's point, listeners overlap to propose a new turn direction. In example (14), the primary speaker A was trying to move on to another topic when B unintentionally interrupted by asking following up questions on the previous topic.

Japanese

(14) A: 続い... ^ん？
tsuzui ... ^n ?
'Moving on to… pardon?'
B: ^漫画だけ？
^manga dake?
'Only comics?'

Although the overlappees were unable to finish their utterances because of the interruption, in order not to cause further competitions and to show politeness on accessing the turn, they tend to yield their turns to the other speaker.

5.3.2 *Pauses* Besides two speakers speaking at the same time, there are also times in CMC conversations where no speakers are willing to take up the turn , hence resulting in different types of silences. These silences often occur after overlaps when both speakers drop out their turns at the same time. In example (15), followed by the overlap, A and B have entered a short silence before B self-selects himself to take up the turn.

Cantonese

(15) A: **咁我就用** (pause)
gam2 ngo5 zau6 jung6
'So I will be using (pause)'
B: **咁你就用** (pause)...negative 呢個題材
gam2 nei5 zau6 jung6 negative ni1 go3 tai4 coi4
'So you will be using (pause) negativity as your topic.'

Example (16) shows how speakers fill in the gap using backchannel responses. After A's comment, B responds with the reactive token "yeah". But since A has shown no intentions to continue his turn, B avoided the potential silences by responding to A's comment with another reactive token before re-taking up the previous turn.

English

(16) A: Such a bummer.
B: Yeah (pause)...yeah...right...Umm...How's
How's your mum doing ?

Similarly, there are situations where speakers fail to fill the gaps with backchannels. In example (17), the informative part ends at the A's utterance, therefore B responds to the statement with the RT "***そう***", noticing that A only respond with RT but not re-taking up the turn, B further utter the Reactive expression "*本当**にそう***" trying to express the intention of passing the turn. As neither of them took up the turn, a 2 second silence has occurred before A self-selects himself again.

Japanese
(17) A: 僕らも できること なら 行こうと 思います
bokura mo dekiru kotonara ikō to omoimasu
'We will also be doing this if we can.'
B: **そう...^** **本当にそう**
*sō ...^* *hontōni sō*
'Yeah...that's right'
A: **^うん** (2 seconds of lapse) ありがとうございます
^ un *Arigatōgozaimasu*
^'Yup' 'Thank you'

6. Discussion

The present study illustrates the change in backchannel productions in terms of frequency and variability. The first research question hypothesized that the number of verbal backchannels will reduce in order not to interrupt or express signs of impatience to the primary speakers. Prior research (Clancy et al., 1996) suggested that Japanese speakers use more verbal backchannels than English and Mandarin speakers, which our findings have confirmed that also occurs in the CMC context in general. CMC speakers have displayed an average higher number of verbal responses, compared to FTF interactions. Even though the weakening effects of CMC are the strongest in Japanese, Japanese speakers have displayed the highest backchannel responses per minute across three languages. Moreover, the data also shows that speakers have also formed longer backchannels, which is reported to be a sign of increased formality (Doherty et al., 1997). RTs remain the major type of verbal backchannels, where in the case of English and Cantonese, the use of RTs are dominant. On the contrary, we predict that listeners will increase their non-verbal backchannels to compensate for the loss in verbal backchannels. There is also a diverse use of non-verbal backchannels shown in our data, including head movements like nodding and body language like gesturing. The results have not fully supported our hypothesis on the backchannel patterns, and thus two implications are made.

Firstly, we believe that the increase in verbal-backchannels is caused by the loss of engagement in the CMC context. Even though CMC and FTF interactions enable speakers to fully see and hear each other, the effectiveness of visual cues might be weakened in the CMC medium. As suggested by Doherty et al. (1997), the communicative commitment of people towards screen-based images has become insensitive due to the prevalence of television (mainly smartphones and laptops in recent days). Therefore, the messages conveyed by visual backchannels might be less effective under this speaker's perception. In CMC media, listeners are more insecure on mutual understanding than FTF interactions, thus speakers generally use more verbal backchannels to engage in the conversation as well as to express their understanding towards the information received. In languages where verbal backchannels are less commonly used (i.e., English and Cantonese), non-verbal backchannels are produced higher in terms of frequency and diversity to achieve attentiveness. However, technical limitations should also be considered in evaluating the effectiveness of information transmission. Our study has presented examples of how frame sizes could restrict the use of full-body languages. It presents a dilemma that we have to encounter: If the person is unaware of the camera, it may cause ambiguity in decoding certain messages; in contrast, if the person is well-aware of the camera, they might intentionally limit their responses mostly to the mid-chest level, which turns out decreasing the naturalness of the whole communication. Lastly, it is suggested that technologies could also distort several conversational acts, where unstable network connections are the most common issue that modern users have. With these technical limitations like lagging, the synchronicity of audio and visual channels in the conversation will be reduced. Not only would it distort the shape of the person's gestures, but the delay in the transmission of responses would also confuse certain situations, hence affecting information transmission.

Our second research question looked into the turn-taking behaviors in CMC contexts. In an organized conversation, participants alternate turns and speak one at a time using different turn-taking strategies. The results

confirmed our hypothesis predicting that turn-taking behaviors will be less-organized. Turn-taking in our CMC data was more disrupted in terms of the increase of verbal backchannels and an overall reduction of resumptive openers, which are used as verbal signals of turn-taking. In addition to the assumption that technical inconsistency could lead to the loss of audio turn-taking signals, it is observed that visual signals are significantly important in turn-taking. Although CMC allows speakers to be visible to each other, similar to visual backchannels, subtle turn-taking signals like eye gazes are often neglected due to the people's insensitivity to screen-based images. Moreover, instead of giving direct eye contact into the eyes, speakers in CMC could only maintain certain "eye contacts" through looking into the camera or the screen. Since it is difficult for speakers to grasp the intention through eye contact and other turn-taking cues, a less unorganized conversation compared to FTF interactions is therefore presented in the CMC medium. Due to unorganized turn-taking behaviors, indications like overlapping and silences were observed. Yet, unlike what we have hypothesized, not all overlaps are disruptive in nature. Cooperative Overlaps are reported to help coordinate the process of the on-going conversations (Trouvain & Truong, 2013), implying that it allows listeners to also engage in maintaining the flow, thus increasing their engagement as well as the naturalness of the conversation.

7. Conclusion

The primary purpose of this study is to conduct a close analysis on the backchannels across Japanese, English, and Cantonese, which provides us with a perspective on how the use of computer-mediated devices could affect the naturalness of a conversation.We expect the backchannel behaviours in informal CMC conversations of these three languages would differ in terms of frequencies, variations and turn-takings. The results have confirmed backchannel responses would be affected by the change in modalities despite the differences in languages. Although distinctions between languages are seen, yet we also see certain similarities in the use of this interactional device as well as turn-taking strategies across languages.

The present study also highlights the communicative efficiency of using CMC as a medium for interactions.Visual information, including both non-verbal backchannels and turn-taking signals, are abundant in FTF interactions. In contrast, even our findings have also reported the use of visual signs, they are not as effective as those in FTF interactions. It is usually the bigger ones (e.g., gesturing) that attract the speaker's attention, while subtle ones (e.g., gazing) are neglected. Since our study has included both quantitative and qualitative aspects , we have not carried out a quantitative comparison of a language community that would further establish the use of backchannels in these languages. Moreover, due to the recent pandemic, more people who are unfamiliar with the use of these platforms have switched to CMCs, which adds the factor of familiarity in evaluating the naturalness of different modalities. In our study, because all the speakers are good users of Instagram lives, the difference in familiarity between speakers was not taken into account. For further research, we suggest a quantitative study of backchannels across modalities. Our findings suggested several variations between FTF interactions and CMC through the use of backchannels. Yet, besides backchannels, there are also other non-semantic elements that can be studied in the future, which the data could be co-operated into future understanding towards language learning. What we hope to achieve through this study is to seek linguistic similarities across languages through these untaught elements, which facilitate the adaptation of new forms of communication.

References:

Clancy, Patricia, A. Thompson, R. Suzuki & Hongyin. Tao (1996). The conversational use of reactive tokens in English, Japanese and Mandarin. *Journal of Pragmatics* 26: 355-387

Cutrone, Pino. (2005). A case study examining backchannels in conversations between Japanese–British dyads. Multilingua - *Journal of Cross-Cultural and Interlanguage Communication*, 24(3), 237–274. doi:10.1515/mult.2005.24.3.237

Doherty-Sneddon, G., Anderson, A., O'malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of experimental psychology: applied*, 3(2), 105.

Duncan Jr., Starkey and Donald W. Fiske. (1977). Face-to-Face Interaction: Research, Methods, and Theories, *New York: John Wiley & Sons.*

Fujimoto, Donna. T. (2009). Listener responses in interaction: A case for abandoning the term backchannel*. Osaka Jogakuin College*

Gardner, Rod. (2001). When listeners talk. *Amsterdam: John Benjamins Publishing Co.*

Hanzawa, Chiemi. (2012) "Listening behaviors in Japanese: Aizuchi and head nod use by native speakers and second language learners." *PhD (Doctor of Philosophy) thesis, University of Iowa*, https://doi.org/10.17077/etd.p4yv50ow

Kock, N., Verville, J., & Garza, V. (2007). Media Naturalness and Online Learning: Findings Supporting Both the Significant- and No-Significant-Difference Perspectives. *Decision Sciences Journal of Innovative Education,* 5(2), 333–355.

Kock, Ned. (2005) "Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward E-communication tools," in *IEEE Transactions on Professional Communication, vol. 48,* no. 2, pp. 117-130, June 2005, doi: 10.1109/TPC.2005.849649.

Kraaz, M., & Bernaisch, T. (2020). Backchannels and the pragmatics of South Asian Englishes. *World Englishes. doi:10.1111/weng.12522*

Lempinen, Juliette. (2020) Finnish upper secondary school students' use of backchannels in an English spoken performance test. *Helsingin yliopisto helda.helsinki.fi*

Maynard, Senko. K. (1997). Analyzing interactional management in native/non-native english conversation: A case of listener response. IRAL, *International Review of Applied Linguistics in Language Teaching,* 35(1), 37.

Mizutani, Nobuko, 1983. Aizuchi to ootoo [Backchannels and responses]. *Osamu Mizutani, ed., Hanashikotoba no hyoogen [Expression in spoken language]. Tokyo: Chikmna Shoboo.*

Passarelli, A., Trinh, M. P., Van Oosten, E. B., & Varley, M. (2020). Can You Hear Me Now? The Influence of Perceived Media Richness on Executive Coaching Relationships. *Academy of Management Proceedings, 2020(1), 13211*

Peters, P., & Wong, D. (2015). Turn management and backchannels. In *Corpus pragmatics: A handbook* (pp. 408-429). Cambridge University Press.

Romiszowski, A. J., and Mason, R.(1996) "Computer- Mediated Communication," in *Handbook of Research for Educational Communications and Technology, D. H. Jonassen (ed.), Simon and Schuster Macmillan, New York.*

Schegloff, Emanuel. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. *Analyzing discourse: Text and talk*, *71*, 93.

Shi, Yuxi (2020) "Reducing uncertainties in virtual consultation: the impact of media naturalness and mental model alignment on patient satisfaction in doctor-patient communication. *Doctoral (PhD) thesis, Memorial University of Newfoundland.*

Tao, Hongyin., & Thompson, Sandra. A. (1991). English backchannels in Mandarin conversations: A case study of superstratum pragmatic "interference." *Journal of Pragmatics,* 16(3), 209–223. doi:10.1016/0378-2166(91)90093-d

Tottie, G. (1991). Conversational style in British and American English: The case of backchannels. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics (pp. 254-273). London and New York: Longman Washington: Georgetown University Press.*

Trouvain, *Jürgen*., & Truong, *Khiet*. P. (2013, August). Exploring sequences of speech and laughter activity using visualisations of conversations*. In Proceedings of the Workshop on Affective Social Speech Signals.*

Wan, J. Y. N. (2018). Functions of frequently used back channels in a corpus of intercultural conversations between Hong Kong Chinese (HKC) and native English speakers (NES). *Journal of Intercultural Communication, 46*.

White, Sheida. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society,* 18(01), 59.

Yngve, Victor. (1970). "On getting a word in edgewise," page 568. Papers from *the Sixth Regional Meeting [of the] Chicago Linguistic Society, 1970*