

Pino Cutrone

# A cross-cultural examination of the backchannel behavior of Japanese and Americans: Considerations for Japanese EFL learners

**Abstract:** Listenership (consisting of backchannel feedback) and its effect on intercultural communication were investigated in 30 dyadic conversations in English between Japanese and American participants. The findings of this study demonstrate several differences in how members of each culture used backchannels in terms of frequency, variability, placement, and function. This study also found evidence supporting the hypothesis that backchannel conventions that are not shared between cultures contribute to negative perceptions across cultures. Thus, the findings of this study support the conclusion that listenership warrants more attention in EFL classes in Japan. Further, toward creating a pedagogical framework, this study also provides EFL professionals with a comprehensive account of the listenership of Japanese EFL speakers. Finally, this study also offers potential insights into the linguistic variation of native English speakers. That is, the negative perceptions that the American native English speakers reported of their Japanese EFL speaking interlocutors' listener responses in this study were not as pronounced as those reported by the British native English speakers in a previous study conducted by the same researcher.

**Keywords:** backchannels, Japanese EFL context, discourse analysis, listenership, pragmatic failure, miscommunication across cultures

---

Pino Cutrone: Nagasaki University, Japan. E-mail: k-pino@nagasaki-u.ac.jp

## 1 Introduction

Despite expending a great deal of energy and resources on English education, Japan has, by most accounts, failed in its attempts to produce competent English speakers (Ellis 1991; Farooq 2005; Helgesen 1987; Hughes 1999; Matsumoto 1994; Okushi 1990; Reesor 2002; Roger 2008; Takanashi 2004; Yano 2001). An important skill involved in oral communication is being able to give effective

listening feedback to one's interlocutor, and this is the specific area of intercultural pragmatics that this paper deals with. It is becoming increasingly clear that what constitutes effective feedback seems open to interpretation, and there is great potential for pragmatic failure and misunderstanding when listening styles differ across cultures. This issue, which has not yet been investigated in great depth, provides the impetus for this research. Hence, the general aims of the study described in this paper are twofold: to expand the scope of empirical work examining conversational behavior across cultures (i.e., between Japanese and Americans) and to determine whether "Listenership," the term McCarthy (2002: 49) coined to describe listener feedback, should receive more attention in EFL classes in Japan and elsewhere.

## 2 Defining listenership

### 2.1 What is a backchannel?

While there currently exist several terms associated with Listenership, the one most commonly used in the literature is *backchannel*, which Yngve (1970: 568) describes as follows:

When two people are engaged in conversation, they generally take turns. . . . In fact, both the person who has the turn and his partner are simultaneously engaged in both speaking and listening. This is because of the existence of what I call the backchannel, over which the person who has the turn receives short messages such as *yes* and *un-huh* without relinquishing the turn.

### 2.2 Backchannels versus turns

In this explanation, Yngve (1970) defines backchannels in terms of where they fit in the turn taking context of conversations. While Sacks, Schegloff, and Jefferson's (1974) seminal work outlining the organization of turn taking is commonly used in the field of conversational analyses, Cutrone (2005) and Maynard (1997) have found Markel's (1975: 190) definition of turn particularly useful for the practical purpose of identifying backchannels in the turn-taking context. Thus, according to Markel (1975), a speaking turn begins when one interlocutor starts solo talking; for every speaking turn there is a concurrent listening turn, which is the behavior of one or more nontalking interlocutors present.

In their studies, Cutrone (2005) and Maynard (1986, 1997) adopted Markel's (1975) definition of turn. In doing so, the position taken in these studies was to identify a brief statement as a backchannel and not a primary turn when it serves only to react to what the primary speaker is saying (listening function) and not to add any new information to the conversation (speaking function). Hence, these studies recognized brief questions such as *Is that so?* or *Really?*, which are formed in terms of requests for clarification, as backchannels. However, a question such as *Why did he go?* was interpreted as a full speaking turn because it serves a speaking function in terms of steering the conversation in a new direction. Further, responses to questions are not considered backchannels. Lastly, researchers have to make decisions regarding how to deal with utterances found between turns at talk, i.e., would such utterances be identified as backchannels or part of a turn at talk? Cutrone (2005) and Maynard (1986, 1997) identified utterances as backchannels only when they occurred immediately after the primary speaker stopped talking (within one second) and were followed by a substantial pause before the next turn at talk starts (exceeding one second). This decision was made because it was felt that these backchannels were produced in response to the primary speaker's speech and they occurred before a substantial turn transitional period starts. For a more in-depth definition of backchannels, and the linguistic environs surrounding their identification in this study, readers are requested to consult the work of Cutrone (2005, 2011).

### 2.3 Backchannels and nonverbal communication

In addition to the multitude of different terms used to describe what Yngve (1970) classified as a backchannel, there exists a great deal of variation even within the term backchannel from one research study to the next depending on the particular definition researchers chose to apply in their analysis (Fujimoto 2007; McCarthy and O'Keeffe 2004). Following the work of several linguists (Fries 1952; Dittman and Llewellyn 1968; Kendon 1967, 1977; Yngve 1970), Oreström (1983) and Hall's (1974) identification of backchannels included mainly minimal utterances such as *uh huh*, *yeah*, *mm*, *I see*, etc. Researchers such as Duncan (1974), Duncan and Fisk (1977), and Maynard (1986, 1987, 1989, 1990, 1997) have broadened this definition of backchannels to include sentence completions, requests for clarification, brief statements and nonverbal responses such as head nods and headshakes. In addition, Brunner (1979), Hattori (1987), and Schenkein (1972) have included other aspects of nonverbal communication such as smiles, laughter, and raised eyebrows as backchannels. As the researcher tends to agree that brief utterances and nonverbal cues by the non-primary speaker are also backchannels

in that they too operate as messages to the primary speaker, this broader definition of backchannels was used in identifying backchannels in an earlier study and will again be used in this study.

## 2.4 Types of backchannels

To show the frequencies of different types of backchannels (i.e., variability), the researcher categorized the verbal backchannels in this study according to three types: simple, compound, and complex. To illustrate this distinction, it is useful to first present Tottie's (1991) classification of backchannels and backchannel items, where the former could consist of one or more of the latter. For example, the sequence *yeah, I see* can function as one backchannel but consist of three backchannel items. Similarly, in my categorization, a simple backchannel such as *yeah* is one that has only one backchannel item. A compound backchannel such as *yeah yeah yeah* is one in which one backchannel item exists but is repeated more than once. A complex backchannel such as *yeah, that's right* consists of multiple and varied backchannel items. To account for nonverbal backchannels, which can occur both simultaneously and independently of the three categories above, Cutrone (2005) added the following categories: simple accompanied by a head nod(s), compound accompanied by a head nod(s), complex accompanied by a head nod(s), isolated head nod, multiple head nods, smile, laughter, raised eyebrows, and two or more nonverbal backchannels occurring simultaneously.

A broader categorical distinction involving listenership can be found in what Stubbe (1998: 259) refers to as “the feedback continuum.” At one end of the continuum is listener feedback that is brief and minimally supportive, while at the opposite end is lengthier feedback that demonstrates a higher degree of involvement in the conversation. In specific terms, Cutrone (2011) describes the former, minimal responses, as any simple (non-word vocalization) and/or nonverbal backchannel occurring in isolation, whereas the latter, extended responses, can be defined as the lengthier verbal listener feedback consisting of multiple and varied words as characterized by complex backchannels, irrespective of nonverbal backchannel accompaniment.

## 2.5 Functions of backchannels

Generally speaking, the functions of backchannels can come under what Sacks, Schegloff, and Jefferson (1974) call recipient design, which refers to the efforts made by the participants in a conversation to adhere to one another's speaking

turns. Maynard (1997) has attempted to sum up the previous work in this area by identifying the following six categories: (1) continuer, (2) understanding, (3) agreement, (4) support and empathy, (5) emotive, and (6) minor additions (see examples of these in Cutrone 2005, 2011). The most common functional type, the continuer, is premised on the turn-taking system and specifically on the non-primary speaker forsaking the opportunity to take a primary speaking turn as proposed by Schegloff (1982). While employing backchannels as a means to avoid speaking may not be exactly what Schegloff (1982: 81) had in mind, several of the Japanese L2 English speaking participants in Cutrone's (2005) study admitted that they often provide backchannels as a way to avoid taking primary speaker-ship, yet still contribute to the conversation. This is an important assertion as it underscores the inextricable link between how much (or little) a person speaks with how frequently (or infrequently) they send backchannels. In the case of Japanese EFL speakers (referred to as JEFs hereafter), several intercultural analyses have shown, predictably, a relative lack of primary speaker incipiency in tandem with the frequent use of backchannels. Further, there is ample evidence to suggest that this overreliance on sending backchannels and the subsequent lack of JEF speaker incipiency (in terms of taking the primary speaking turn and asking their interlocutor questions to drive the conversation forward) are perceived negatively in ELF (English as a Lingua Franca) conversations across cultures (Cutrone 2005; Sato 2011). Thus, as it relates to backchannel behavior, this study also examines how much participants spoke in the intercultural conversations, as well as how many questions they asked their interlocutor.

### 3 Listenership across cultures

A number of studies have examined various aspects of listenership across cultures. While certain descriptive areas of listenership, such as frequency, have received a great deal of research attention, other areas, namely variability, function, and perceptions across cultures have been largely neglected. The purpose of this section is to provide a brief but coherent picture of what is known about this phenomenon relative to the Japanese and to, thus, demonstrate where further inquiry would be helpful.

#### 3.1 Frequency

The results from a number of intercultural analyses (Cutrone 2005; Clancy et al. 1996; Crawford 2003; Ike 2010; Maynard 1986, 1990, 1997; White 1989) have

consistently shown JEFs producing backchannels far more frequently than native-English-speaker (NES) interlocutors (i.e., Britons, Americans and Austrians) in both their L1 and L2 English.

### 3.2 Variability

In their intercultural analyses, Maynard (1997) and Cutrone (2005) found that the JEFs' backchannels consisted mainly of non-word vocalizations and head nods, while the NESs (American and British respectively) displayed greater variability in the types of backchannels they sent.

### 3.3 Discourse contexts

While various studies have provided different degrees of detail in their analyses, a general finding has been that Japanese discourse contexts favoring backchannels varied considerably (in both the L1 and L2 English), while grammatical completion points (i.e., clausal boundaries), especially coinciding with a pause, were the single most important discourse contexts for NESs (Cutrone 2005; Maynard 1986, 1990, 1997; White 1989).

### 3.4 Simultaneous speech

A common finding in the research is that Japanese people, regardless of whether they are speaking English or Japanese, tend to backchannel more frequently than NESs, and a great portion of these backchannels are sent during the primary speakers' speech, creating simultaneous speech (Cutrone 2005; Hayashi 1988; Maynard 1997). As various researchers such as Lebra (1976) and Mizutani (1982) have asserted that some may take these frequent interjections as a sign of the listener's impatience and demand for a quick completion of the statement, White (1989) and Cutrone (2005) conducted correlational analyses to find out what effect Japanese people's frequent backchannels might have on their cross-cultural interlocutors perceptions of them. Interestingly, the 10 Americans in White's (1989) study perceived more frequent backchannels sent by their Japanese interlocutors as a positive trait (i.e., showing signs of comprehension, encouragement, and interest and concern), while the eight Britons in Cutrone's (2005) intercultural analysis perceived more frequent backchannels as interruptions and signs of impatience.

### 3.5 Function

While this has been an under-researched area overall, there may be some key functional differences across cultures as Blanche (1987) and Cutrone (2005) have provided anecdotal evidence of Japanese people providing unconventional backchannels in English, such as by employing continuer, understanding, agreement, and/or empathy/support type backchannels in situations when they did not understand or disagreed with what their interlocutor was saying at the time. Further, several studies involving the intercultural analyses of communication styles have shown that the Japanese L2 English speakers in these studies spoke less than NESs, did not elaborate as much, and were less likely to engage in small talk (Cutrone 2005; Hill 1990; Sato 2008). This is potentially a source of misunderstanding in an English conversation as the importance of making small talk, taking the initiative to speak, and elaboration toward making a positive impression have been documented by several sources (Cutrone 2005; McCarthy 2003; McCroskey 1992; Ross 1998; Sato 2008; Stubbe 1998; Yashima 2002).

## 4 Research questions

Research into the area of backchannel behavior, and particularly its effect on intercultural communication (IC), is in its infancy and, thus, many questions remain. Most of the research to date has been descriptive and has focused on detailing patterns of the backchannel output of various groups in terms of frequency and discourse contexts, and to a much lesser degree, variability, and function. Research Question (RQ) 1 thus aims to corroborate (and add to) the findings of previous studies involving IC among JEFs and proficient speakers of English. Considering an area of backchannel behavior only previously touched upon by White (1989) and subsequently re-examined in greater depth by Cutrone (2005), RQ 2 seeks to add a qualitative element to the study of listenership by examining the extent to which differences in backchannel conventions across cultures affect IC.

In light of the research to date, the following RQs have been formulated.

- RQ 1: How do participants from each culture, Japanese and American, use backchannels differently in English in terms of frequency, variability, discourse contexts favoring backchannels and concerning function?
- RQ 2: Will the differences in backchannel behavior across cultures affect communication and/or lead to miscommunication, and if so, how will this occur?

In attempting to bridge the gap between research and practice, the answers to these questions will go a long way toward helping EFL professionals establish a pedagogical framework for the inclusion of listenership in the EFL curriculum in Japan.

## 5 Methodology

As the research questions are complex and multifaceted in nature, various methodological frameworks were considered in conducting this study. Ultimately, it was decided that the most reliable and valid approach in collecting data involved a combination of the three methods being considered: observations, questionnaires, and interviews. The following subsections will describe the process by which participants were selected, the procedures for collecting data, and the methods used for analyzing data.

### 5.1 Participant selection

Overall, this study involved 43 participants. Concerning the conversational aspect of this study, this study used 30 JEFs born and raised in Japan and three L1 American English speakers from the United States. The gender division of the participants was 24 females and six males for the Japanese group, and one female and two males for the American group. At the time of the study, all participants resided in Nagasaki Prefecture in Japan and were university students ranging in age from 18–20. Additionally, 10 American participants (i.e., NES observers in this study), who, unlike the American exchange students participating in the conversations, did not participate in the conversations and/or have any particular affiliation to Japan, were used to provide unbiased assessments of the JEFs in the conversations. These participants ranged in age from 22–48 and were from various parts of the United States. Participating of their free will and understanding the nature of the study, all 43 participants read and signed a Participation Consent Form and were given explicit instructions regarding this study and their role in it. All forms were typed in English with Japanese translations provided to the native Japanese speakers in this study to ensure these participants had a full understanding of the contents in each form. In referring to participants in this study, pseudonyms have been given to protect participants' privacy. For the sake of mnemonicity, Japanese names were given to the JEFs and English names were given to the NESs.



Concerning the JEFs, there was some variation in their English proficiency level as 17 of the participants were within the 500–650 TOEIC range, 10 participants ranged from 350–499 and three participants ranged from 651–700. The Japanese participants and their parents were required to have been born in Japan. Also, it was required that these participants have only limited experiences abroad. This was judged to be less than 100 days throughout the course of their lives, and never for a period of more than 20 days at one time.

Regarding the NESs in this study, they all spoke American English and had parents who were born in the United States. Further, they had all lived in Japan for less than a year. A short period of residence in Japan was thought to be important because the longer the contact with members of the other culture, the greater the possibility of changes in backchannel use and interpretations (Locastro 1987; White 1989).

Moreover, as part of this study involved observing intercultural dyadic conversations, the range of participants selected made it possible to create intercultural dyads in which both participants contained common sociolinguistic characteristics where gender and age were concerned. All dyads were paired according to gender, and in all cases, the ages of the participants in each dyad were within two years of each other. Another concern was the extent to which the participants in each dyad knew each other. Taking into account the participants who were available, this study allowed for some participant contact prior to the study but made every effort to pair dyads together who were less familiar with each other.

## 5.2 Procedures in collecting data

### 5.2.1 Observations

Following (and in some cases supplementing) the procedures used in Cutrone's (2005) earlier study, the researcher employed three methods of collecting data for this study consisting of observations, questionnaires, and interviews. The first of these three methods, observations, involved the video recording of 30 intercultural dyadic conversations in English between JEF and NES participants. The conversations were conducted in the researcher's office. Video recording equipment consisted of a Sony digital video camera (Model No. DCRTRV30), which was placed unobtrusively in the corner of the room (no additional microphones were used). At the time of the video recording, only the participants were present in the room when the conversation was taking place. Each of the 30 dyads was instructed to talk as casually and as naturally as possible about

anything they like for a period of 30 minutes, of which only the middle three minutes of each conversation were included as data to be transcribed. Following the work of White (1989), Maynard (1990, 1997) and Cutrone (2005), it was thought that the participants would become less conscious of the camera as the conversation progressed, and that the middle part of the conversation would be the most natural as it avoids the awkwardness which often occurs at the beginning and end of conversations between people who do not know each other well.

### 5.2.2 Questionnaires

Following Cutrone (2005) and White (1989), the questionnaires used in this study were modified versions of the one Hecht (1978) designed to measure conversational satisfaction. The questionnaire given to the conversational participants consisted of a 15-item inventory. Questions were closed-ended, consisting of statements on a Likert-scale ranging from one to seven. The researcher modified the questionnaire from Hecht's (1978) original and the one White (1989) used because pilot studies revealed that some vocabulary and some of the statements, which contained double negatives, confused participants. The first set of questionnaires was given to each participant in the dyad directly after their video recorded conversation and simultaneously completed in separate rooms. Subsequently, after all the conversations (and interviews) were completed, all 30 of the video recorded conversations were copied to Digital Video Discs and given to a group of American observers (in random order) to assess their perceptions of the JEFLs' performances. To this end, the NES observers completed another version of Hecht's questionnaire upon watching each video recorded conversation.

### 5.2.3 Interviews

Taking place directly after the participants of the intercultural dyadic conversations completed their conversations and subsequent questionnaires, the interviews involved the two members of each dyad being interviewed separately and in succession, with the NES first and the JEFL second. No one else was present at the time of the interview, as the researcher asked the participant not being interviewed to wait in another room while the interview was being conducted. None of the interviews took longer than 20 minutes. The interviews consisted of the researcher playing back a portion of the video recorded conversation and asking

each participant a few questions pertaining to the listening behavior displayed in the conversation. The researcher took field notes and audio recorded all interviews (using a Sony digital voice recorder; Model No. ICDUX5122). The interviews were semi-structured in that the researcher had a general plan but did not enter with a predetermined set of questions, as some of the questions were guided by the circumstances in the video recorded conversations and the responses of the interviewee. In an effort to help participants feel more comfortable, questions were sequenced to begin with general inquiries and gradually move toward more specific and potentially sensitive questions.

The primary aim in the interviews with the American participants was to learn how they perceived their Japanese interlocutors' listenership. A major part of the interview involved playing back the video recording and asking the American interviewee to comment on the listening behavior of their Japanese interlocutor. In instances singled out for analysis (where some of the JEFs' backchannels occur), the researcher stopped the video recording and asked specific questions such as *What function do you think that head nod serves?*, *Do you think s/he understands what you are saying here?*, and follow-up questions such as *Why do you think so?* The researcher made a note of any data thought to be useful in the subsequent interview with the Japanese member of the dyad.

In interviewing the Japanese participant, the researcher had two main objectives: to gain insights into why JEFs use backchannels the way they do and to determine if there were any misunderstandings or miscommunications caused by their use of backchannels in the video recorded conversations. Regarding the latter, this involved the interviewer asking the JEFs to comment on what they were feeling at certain points in the conversation, as well as the intended functions of their backchannel responses. In particular, the researcher sought to examine what the JEFs were doing when they did not understand and if they were indeed feigning understanding as various researchers allege is common for Japanese backchannel behavior (see Section 3.5). In the successive interviews, the researcher documented any notable differences between the Japanese participants' backchannel intentions and their NES interlocutors' perceptions on a data record sheet. In the cases where the Japanese participants' backchannel explanation differed greatly from their NES interlocutor's interpretation, the interviewer asked potentially sensitive follow-up questions such as *Why did you nod if you did not understand what he/she was saying here?* and *Why did you say yeah if you disagreed with what he/she was saying?* In cases where the interviewee seemed uncomfortable in answering, the interviewer did not persist with this line of questioning and instead shifted to a less sensitive area.

## 5.3 Data analysis

This section describes how the data produced by observations, questionnaires, and interviews are analyzed in this study. First, it is necessary to point out the obvious limitation that the Japanese group consisted of 30 participants making up 30 sets of data, whereas the 30 sets of data corresponding to the American interlocutor group were generated by only three American participants, who each participated in multiple conversations. The reasons for this asymmetry are three fold: (1) due to practical constraints, the number of NES exchange students available to the researcher were limited to the three participants in the study; (2) following the researcher's earlier work in this area, this study modestly attempts to confirm or refute previous findings in this area, as well as provide a launching pad for more in-depth inquiries into this area of pragmatics; (3) lastly, focusing on the 30 JEFLs, the cross-sectional study described herein was part of a longitudinal study designed to determine the effects of pedagogical interventions on JEFL learners' backchannel behavior over time (Cutrone, pending). Hence, in an attempt to draw attention to the real-world problems experienced by Japanese L2 English speakers where backchannel behavior is concerned, the ultimate goal of this study into intercultural pragmatics is to set the stage for pedagogical applications.

Due to the basic violation of independence between the conversations described above (i.e., the NESs participated in multiple conversations), it is not appropriate to apply inferential statistics to various aspects of the data and, thus, only descriptive statistics are used to illustrate many of the differences between two cultural groups. However, the data collected from the 10 American observers' questionnaires did not violate the basic assumption of independence, and, thus, inferential statistics can be applied to this aspect of the data analysis. Concerning inferential statistics, both parametric and nonparametric statistical tests have been applied depending on the type of data analyzed. Using the Statistical Package for the Social Sciences (SPSS), version 14.0, the researcher has used two-tailed tests and set alpha levels ( $\alpha$ ) at 0.05.

### 5.3.1 Observations

The first step in analyzing the data was to transcribe the 30 conversations according to the linguistic environs outlined in Section 2. To strengthen the internal reliability of this analysis, the researcher transcribed the conversations with the assistance of a colleague. This involved the colleague double-checking the re-

searcher's initial transcriptions to ensure that the conventions showing the features of language were accurately presented in the transcriptions.

### 5.3.2 Questionnaires

The first set of questionnaires was designed to convey how conversational participants from each culture generally felt about one another after the conversations. The second set of questionnaires demonstrates how the American observers perceived the JEFs' listening behavior and whether there was any correlation between the American observers' ratings and the observed behaviors of the JEFs' listenership.

### 5.3.3 Interviews

Supplementing the data provided by the questionnaires, the interviews also seek to shed light on perceptions across cultures. To this end, the texts of the interviews were transcribed verbatim and then investigated for emergent themes. These themes were based on two considerations: the various patterns found in the interviews, and in answer to questions pertaining to key concepts in this study. Addressing the former, analyses of the interview text involved synthesizing the data in such a way as to produce possible patterns, yet without misrepresenting or distorting the data. Concerning the latter, analyses of the interview text also focused on finding answers to questions which delve deeper into the reasons JEFs use backchannels the way they do.

## 6 Quantitative results

### 6.1 Frequency

The conversational data showed that the JEFs provided considerably more backchannels than the NESs. Table 1 details the observed frequency of backchannels as they pertain to each culture and also reports on the willingness each group had to speak and carry the conversation in terms of the amount of words each group uttered. It is clear to see that the American group spoke a great deal more averaging 263.77 words per conversation while the Japanese group averaged 62.43 words.

**Table 1:** Differences in frequency of backchannels across cultures

JEFs = 30 NESs = 3	Total Backchannels		Total Words		Average number of interlocutor's words between backchannels	
	Japanese	Americans	Japanese	Americans	Japanese	Americans
Overall Total	1265	289	1873	7913	–	–
Mean ( $\bar{x}$ )	42.16	9.63	62.43	263.77	6.46	7.38
SD	13.18	6.11	31.39	247	1.36	2.12
Median	39.5	8	56.5	63.68	6.31	6.92
Range	47	289	114	192	5.5	7.78

As Section 5.3.2 touched upon, Spearman rho correlational analyses were conducted to determine whether there was a possible association between the frequency of a particular feature of the JEFs' listenership and the NES observers' responses of each item on the conversational satisfaction questionnaire (see Appendix). A statistically significant correlation was found between the number of backchannels the JEFs sent and the NES observers' ratings on Item 2 ( $p < 0.029$ ), Item 5 ( $p < 0.023$ ) and Item 6 ( $p < 0.003$ ). Thus, as the number of JEF backchannels increased, (Item 2) the more the NES observers believed the JEFs to be showing understanding, (Item 5) the more NES observers felt the conversation went smoothly and (Item 6) the more the NES observers believed the JEFs to be encouraging their interlocutor to continue speaking. A statistically significant relationship was found between the number of words the JEFs uttered and the NES observers' ratings on Items 1, 2, 7, 15, and 16 (see Appendix). The direction of the significant correlation for all these items shows that greater word outputs were associated with positive ratings.

Table 2 shows that the NESs made a much greater effort to ask questions in the conversations as they averaged 2.33 questions per conversation as compared to 1.27 questions posed by the JEFs.

**Table 2:** Number of questions across cultures

JEFs = 30 NESs = 3	Japanese				Americans			
	Total	$\bar{x}$	SD	Median	Total	$\bar{x}$	SD	Median
# of Questions	38	1.27	.87	1	70	2.33	1.88	2

## 6.2 Variability of use

Tables 3 and 4 distinguish the frequency of specific sub-types of verbal and non-verbal backchannels across cultures respectively.

**Table 3:** Types of verbal backchannels across cultures

Type	Japanese Backchannels N = 30				American Backchannels N = 3			
	Total	$\bar{x}$ % of Total BCs	SD	Median	Total	$\bar{x}$ % of Total BCs	SD	Median
Simple (isolated)	147	13.21	9.27	11.51	25	7.81	12.96	0
Simple with head nod(s)	476	35.71	16.77	37.85	112	37.49	34.06	35.42
Compound (isolated)	14	1.03	1.84	0	5	1.77	4.8	0
Compound with head nod(s)	39	2.74	2.18	2.74	15	4.55	8.53	0
Complex (isolated)	31	2.79	4.46	1.96	22	11.15	13.98	7.14
Complex with head nod(s)	68	5.11	4.05	4.44	50	20.83	17.08	16.67
Japanese (all complex)	11	1.11	4.56	0	0	0	0	0

Some of these categories combine to form two separate and distinct superordinate classes of backchannel responses: extended responses and minimal responses (see Section 2.4). As Table 5 illustrates, the mean percentage of backchannels constituted by minimal responses was much greater for the JEFs (JEFs: 75.96%, NESs: 62.88%), whereas that constituted by extended responses was much greater for the NESs (NESs: 35.38%, JEFs: 7.95%).

Using Spearman rho, a strong statistical significant relationship (at the 0.01 level) was found between the number of extended responses the JEFs sent and the NES observers' ratings on items 1, 2, 4, 6, 7, 11, 13, 15, and 17, while a significant relationship (at the 0.05 level) was observed for items 5 and 16 (see Appendix). The direction of the significant correlation for all these items shows that more frequent extended responses were associated with more positive ratings.

As shown in Table 6, backchannels as repetitions (REPS) are distinguished by two types: clarifications (CLARS) – with a rising intonation, and non-clarifications (NON CLARS) – with a steady or falling intonation. The mean percentages of backchannels constituted by repetitions were fairly similar overall (JEFs: 3.3%, NESs: 2.19%); however, the NESs employed a greater proportion of backchannels constituted by clarification (JEFs: .5%, NESs: 1.64%), while the

**Table 4:** Types of nonverbal backchannels across cultures

Type	Japanese Backchannels N = 30				American Backchannels N = 3			
	Total	$\bar{x}$ % of Total BCs	SD	Median	Total	$\bar{x}$ % of Total BCs	SD	Median
Total nonverbal BCs (isolated)	479	36.79	17.98	34.52	60	19.45	18.10	17.42
Single head nod (isolated)	222	15.75	12.4	16.33	36	12.25	16.39	2.38
Multiple head nods (isolated)	163	10.75	11.02	7.33	17	4.37	6.25	0
Smile (isolated)	0	0	0	0	1	.56	3.04	0
Smile with verbal backchannel	31	2.38	3.7	.82	4	2.26	7.6	0
Laughter (isolated)	93	8.31	6.86	6.11	4	1.07	3.37	0
Laughter + verbal backchannel	38	2.6	4.2	0	3	1.55	5.43	0
Raised Eyebrows (isolated)	0	0	0	0	1	.24	1.3	0
Raised Eyebrows + verbal BC	19	1.53	2.39	0	4	1.15	3.26	0
≥ 2 nonverbal BCs occurring simultaneously (isolated)	1	.08	.45	0	1	.29	1.1	0
≥ 2 nonverbal BCs occurring simultaneously with verbal BC	6	.5	1.62	0	2	.32	1.74	0

**Table 5:** Minimal responses versus extended responses across cultures

Type of Backchannel	Japanese Backchannels N = 30				American Backchannels N = 3			
	Total	$\bar{x}$ % of Total BCs	SD	Median	Total	$\bar{x}$ % of Total BCs	SD	Median
Minimal Responses	1008	75.96	17.52	79.77	190	62.88	35.97	66.67
Extended Responses	99	7.95	5.59	8.39	72	35.38	30.53	32.58



JEFLs sent a greater proportion of backchannels constituted by non-clarification (JEFLs: 2.82%, NESs: 0.56%).

**Table 6:** Use of repetitions as backchannels across cultures

	Total REPS	$\bar{x}$ % of Total BCs	SD	Median (M)	Total CLARS	$\bar{x}$ % of Total BCs	SD	M	Total NON CLARS	$\bar{x}$ % of Total BCs	SD	M
Japanese	40	3.3	1.59	3.16	6	.5	1.06	0	34	2.82	1.54	2.87
Americans	5	2.19	1.06	0	4	1.64	5.1	0	1	.56	3.04	0

### 6.3 Discourse contexts favoring backchannels

Table 7 reports two main statistics regarding the discourse contexts of each group's backchannels (BCs): (1) the mean percentage of opportunities (Opps) that each discourse context attracted backchannels (with SDs), and (2) the mean percentage of backchannels constituted by each discourse context (with SDs). Notably, the results show that the completion of a grammatical clause, and particularly one accompanied by a pause, provides the single most significant contextual cue for backchannels for both cultural groups (JEFLs: 63.87%, NESs: 56.8%).

### 6.4 Simultaneous speech backchannels (SSBs)

Table 8 illustrates the great discrepancy between the two groups: JEFLs provided 242 backchannels during the primary speaker's turn at talk, while the NESs produced only 29. When the results are reformulated to take into account the total number of backchannels uttered by each group, the difference between the two groups has decreased but still remains substantial as 18.1% of the JEFLs' backchannels were uttered while their interlocutor was taking a turn at talk and only 10.77% of the NESs' backchannels were provided in this context.

Using Spearman rho, a strong significant relationship (at the 0.01 level) was found between the number of non-laughter SSBs the JEFLs sent and the NES observers' ratings on items 1, 2, 4, 5, 6, 7, 11, 13, and 17, while a significant relationship (at the 0.05 level) was observed for items 3, 8, 12, 15, and 16 (see Appendix). The direction of the significant correlation for all these items indicates that more frequent non-laughter SSBs were associated with more positive ratings.

**Table 7:** Discourse contexts of backchannels across cultures

Discourse Contexts	Japanese N = 30			Americans N = 3		
	Total	$\bar{x}$ % of Opps (SD)	$\bar{x}$ % of BCs (SD)	Total	$\bar{x}$ % of Opps (SD)	$\bar{x}$ % of BCs (SD)
At or directly after a pause	576	42.87 (12.95)	46.56 (18.69)	150	23.37 (9.91)	55.72 (23.07)
<b>Clausal Boundaries:</b>						
At or near any CB	779	60.13 (12.19)	60.26 (14.43)	121	43.03 (22.56)	37.87 (21.18)
At or near internal CB	42	27.33 (30.13)	2.88 (2.01)	7	11.67 (28.77)	1.69 (4.17)
At or near final CB	733	61.83 (15.03)	56.87 (14.26)	110	46.57 (24.28)	35.07 (20.86)
<b>At or directly after points in which CBs and pauses occur together</b>						
At or directly after any CB accompanied by a pause	419	62.03 (16.16)	33.64 (13.65)	72	53.13 (33.05)	23.95 (18.86)
At or directly after internal CB accompanied by a pause	19	26 (38.18)	1.4 (2.5)	2	6.67 (25.37)	.34 (1.29)
At or directly after final CB accompanied by a pause	415	63.87 (16.12)	33.64 (13.75)	71	56.8 (32.59)	23.76 (18.9)
<b>Gesticulation:</b>						
After primary speaker's nonverbal gesture	79	28.5 (20.21)	7.13 (7.64)	31	20.57 (19.89)	12.59 (14.2)
After primary speaker's head nod(s)	104	42.73 (22.52)	8.94 (7.18)	27	22 (24.02)	14.34 (21.05)

**Table 8:** SSBs across cultures

SSB Types	Japanese N = 30			Americans N = 3		
	Total	$\bar{x}$ % of BCs	SD	Total	$\bar{x}$ % of BCs	SD
Non-laughter	189	13.99	8.56	25	9.21	12.61
Laughs	53	4.11	5.18	4	1.56	4.2
Total SSBs	242	18.1	10.98	29	10.77	12.74

## 6.5 Questionnaires: Examining participants' conversational satisfaction

Overall, as shown in Table 9, the results from the questionnaire were comparatively similar across cultures, and indicated that participants generally had positive impressions from their conversations with their cross-cultural interlocutors. Nonetheless, a few notable exceptions in which the JEFs expressed substantially less satisfaction ( $>1$ ) were found in Items 2, 3, 5, and 6. First, concerning Item 2, the JEFs' score of 4.2 demonstrates that they felt that they might not have represented themselves fairly in the conversations. The NES interlocutors' score of 1.87 in Item 2, conversely, indicates the general perception that their JEF counterparts expressed themselves in a way that seemed sincere and authentic. This considerable difference between the groups' scores (2.33) in corresponding items would seem to highlight a gap in the IC that took place. Moreover, the JEFs were markedly less convinced that the NESs were understanding them (difference of 1 in Item 3), were interested in what they had to say (difference of 1.3 in Item 5), and that the conversation went smoothly (difference of 1.95 in Item 6).

# 7 Qualitative results

## 7.1 Factors affecting listenership

While the main focus of the interviews was to gain insights into participants' listening behavior, this initial section reports on some of the experimental and/or design issues that may have influenced participants' behavior in the conversations. To this end, the researcher sought to determine the extent that participants may have changed the way they behaved because the conversations were being video recorded (known as Observer's Paradox), as well as the degree to which the NESs accommodated to their intercultural interlocutors. Concerning the presence of Observer's Paradox in this aspect of the study, the overwhelming majority of the participants (all three of the NESs and 27 of the 30 JEF participants) indicated that they were not conscious of the camera once the conversation developed. Further, considering the extent to which the NESs accommodated to their intercultural interlocutors, all three NESs interviewed in this phase of the study acknowledged that they did indeed modify their behavior and speech in various ways to better communicate with their JEF interlocutors, which included slowing down their speech, initiating comprehension checks and carefully choosing

**Table 9:** Participants' conversational satisfaction

Items on the Questionnaire	Japanese N = 30	American N = 3
N.B. Scale ranges from 1 (strongly agree) to 7 (strongly disagree).	$\bar{x}$ (SD)	$\bar{x}$ (SD)
1. S/he let me know that I was communicating effectively.	2.27 (1.05)	1.68 (.86)
2. I felt I was able to present myself fairly during the conversation (for JEFs only).	4.2 (1.57)	–
2. The feelings that my partner expressed by means of listening feedback during the conversation seemed authentic (for NESs only).	–	1.87 (1.49)
3. S/he showed me that s/he understood what I said.	2.4 (1.1)	1.4 (.86)
4. S/he showed me that s/he listened attentively to what I said.	1.33 (.71)	1.3 (.84)
5. S/he expressed a lot of interest in what I had to say.	2.67 (.88)	1.37 (.96)
6. The conversation went smoothly.	3.53 (1.48)	1.58 (1.12)
7. S/he encouraged me to continue talking.	1.37 (.56)	1.43 (1.17)
8. S/he seemed impatient.	6.53 (1.2)	6.73 (6.9)
9. S/he seemed cold and unfriendly.	6.97 (.18)	6.83 (.59)
10. S/he was polite.	1.57 (1.55)	1.1 (.31)
11. S/he appeared warm and friendly.	1.17 (.38)	1.77 (1.05)
12. S/he was impolite.	6.83 (.91)	6.93 (.25)
13. S/he appeared interested and concerned.	1.97 (1.3)	1.4 (1.07)

Table 9 (cont.)

Items on the Questionnaire	Japanese N = 30	American N = 3
14. S/he interrupted me.	6.83 (.59)	6.53 (1.01)
15. My conversation partner seemed to want to avoid speaking.	6.27 (1.34)	6.53 (1.25)

level-appropriate vocabulary. One major theme to emerge from this line of questioning is that all NESs admitted that they generally have much more patience and tolerance for misunderstandings in IC than they would have communicating with other NESs. The data also showed that some of the JEFL participants also seem to modify their behavior when speaking across cultures, albeit in a different way. Many of the JEFL respondents indicated that, while they would like to accommodate toward the NESs, their behavior and performance in IC encounters is mostly affected by the anxiety they feel, which stems from not knowing how to act in these situations.

## 7.2 Functions of backchannels across cultures

Conveying that the functions of backchannels were largely determined by the context of the conversation at that time, participants from both cultures cited the continuer type as the one they used most frequently and also commonly employed backchannels to show comprehension, to show agreement, and to show empathy and support. Some functional differences between the two groups began to emerge, however, when participants were asked to explain the thought processes underpinning some of these common backchannels. While the three NESs described their continuer backchannels in terms generally associated with Schegloff's (1982) original definition (i.e., to let the primary speaker continue their turn), all 30 of the JEFLs reported that they commonly use backchannels as a means to ensure continued participation and inclusion in the conversation, and 19 of them acknowledged that they also sometimes send backchannels as a way to avoid speaking due to shyness, lack of confidence in their English ability, and not knowing what to say next. These feelings are expressed in Madora and Keiko's responses below:

Madora: I very much want to speak, but the conversation is too fast, and I am too shy I think. So then, I just give *aizuchi* (the Japanese word for back-channel) to be a part of it.

Keiko: It's difficult to think what I want to say, so I sometimes just nod and say yeah yeah.

Meo and Taro's responses below point directly to Japanese culture as the basis for this type of behavior:

Meo: I think foreigners love speaking and attention and Japanese are very shy and like mostly listening and not speaking, so we give lots of *aizuchi*. It's Japanese nature, I believe.

Taro: I think Japanese try to be polite and let the other person talk. We give many *aizuchi* to support the other person when they speak.

This study also observed notable cultural differences in the types of backchannels used when participants did not understand or disagreed with what their interlocutor was saying in the intercultural conversations. The NESs did not employ any of the frequently used backchannel types (i.e., continuer, understanding, support and empathy, and agreement) during or after the 11 instances they reported of not understanding and the 1 instance of disagreement, whereas the JEFs produced these types of unconventional backchannels 98 times during or after the 139 instances (71%) they reported not understanding what their interlocutor was saying. Regarding the two instances the JEFs disagreed with their interlocutor, they provided continuer type backchannels both times (100%).

Explaining why they had employed continuer, understanding, agreement, or empathy/support type backchannels in situations when they did not understand and or disagreed with what their interlocutor was saying at the time, some of the JEFs again mentioned cultural factors and issues concerning politeness. That is, as the following responses demonstrate, many JEFs seemed to place a far greater priority on preserving harmony and saving face rather than expressing their opinions clearly and honestly.

Yukiko: I don't like to argue with people, so I always agree with their saying.

Akie: I think interrupt partner is no good even I disagree. Nice communication is most important for us Japanese. I don't want partner to lose the face.

Shio: Sometimes I can't understand anything, but I want to let my partner talk more. It's Japanese way to not cut (interrupt).

Kouki: It is rude to interrupt, so I fake understanding sometimes.

Madora: Even I don't understand, I nod and give *aizuchi* because don't want to lose nice atmosphere.

Some JEFs commented that they sometimes employed backchannels when they did not understand to avoid embarrassment and humility. This can be seen in the following excerpts:

- Mikki: I am not confident and so ashamed (at) my bad English. I pretend that I can understand so partner will not lose respect for me.
- Takanori: I am scared if I don't understand my partner will think I (am) stupid, so I just say *un un* and nod always, even I don't know his saying.
- Rika: Japanese often don't show true feeling because they want to make nice feeling. Another reason, maybe, is because we are afraid that we look a little *baka* (stupid), so we just pretend to understand, and, of course, it's more polite this way.

Lastly, as shown in the following responses, some JEFs indicated that they sometimes employ backchannels as a means to allow themselves a few extra seconds to process information and decide what they will say next.

- Yoko: It's hard to keep up with fast conversation, so I sometimes just nod or say *un un* to give myself some moment (time) to think.
- Hiro: I need time to translate in my head, so I give *aizuchi* for time and to keep partner talking until I am ready.
- Yuki: Sometimes, I can understand enough but cannot get the words out of my mouth fast enough. I think giving *aizuchi* is good to let me think what I want to speak next.

All three NESs responded that they did not use backchannels to gain additional processing time, or at least were not aware of it if they did.

### 7.3 Perceptions of backchannels across cultures

As the results of the conversational satisfaction showed (see Section 6.5), both groups had favorable impressions overall of their cross-cultural interlocutors after their intercultural conversations. However, in noting that the NESs' listener responses were much more exuberant and variable than the JEFs', both groups appeared to have different notions as to how greatly this might affect IC. As the following responses demonstrate, the JEFs recognized the difference in

exuberance and variability; however, they did not necessarily convey that it would be harmful to IC:

- Kazuya: Japanese often say the same kind of *aizuchi*; not like Americans, they say so many different things and give big and livelier reactions.
- Michiko: My partner's reactions were much bigger than mine. I am not so comfortable giving such big ones as that.
- Aria: Japanese give usually the small *aizuchi*. Americans are more big um how do you say *nigiyaka* ('lively') in English?

Conversely, as the following extracts show, the NESs all mentioned the JEFs' lack of variability and exuberance as a source that might negatively affect IC.

- Betty: It was difficult to tell if a few of them were really listening and getting it cos, ya know, all the reactions were kinda similar . . . I wanted more enthusiasm I guess.
- Jerry: I know it's tough speaking a second language, but I wish their reactions would better fit the situation. For instance, if I told them I had won a million dollars, I want to hear *wow* or *that's great* rather than just see a little head nod or hear a little grunt.
- Jason: Many of the reactions didn't feel very natural. I kind of got the impression that he may have been just going through the motions. I wasn't even sure if he understood what I was saying because all the responses were the same.

Related to Jason's response directly above, the NESs also voiced confusion and/or minor irritation when they detected that the JEFs employed backchannels in ways that seemed unconventional in English such as when they do not understand. Betty and Jerry's responses also demonstrate this:

- Betty: I sometimes feel that my partner was just nodding and laughing to be polite whether she understood or not. Americans would never do that.
- Jerry: I have a hard time telling if Japanese people are really understanding me. For instance, when my partner said *mhm* here (points at the screen), I still don't know if he knows what character I was talking about. This always used to happen when I first got here a few months ago. People would just nod away even though they haven't faintest clue what I'm saying.



Regarding the more frequent backchanneling style of the JEFL group, the NESs' responses were mixed. Whereas one NES was not conscious of any disparity in backchannel frequency, the other two NESs had differing opinions. That is, Betty seemed to enjoy the rather large amount of listener feedback, while Jerry believed it to be excessive and interruptive, particularly while he was speaking.

**Betty:** My partner nodded a lot and, generally, I like getting a lot of responses when I'm talking. I couldn't tell if she understood everything, but at least she was trying. For me, the more (responses), the merrier.

**Jerry:** It was difficult to know if he wanted to talk or what. I mean sometimes he would interrupt me with grunts or words I had just said, and I'd try to let him talk and he wouldn't take it any further and we'd have these awkward silences. If that was supposed to be listening feedback, then it felt a little strange to me. Laughs are fine, but it was the other stuff that I didn't really get. I guess I prefer the feedback after I'm done speaking and not during (my speech).

Lastly, although the NESs generally accepted (and expected) that they would have to carry the conversation in these NS-NNS exchanges, they admitted to wanting their JEFL interlocutors to be more involved in the conversations. In some instances, the JEFLs' perceived lack of involvement seems to have caused the NESs to question how interested the JEFLs were in the conversation. This lack of JEFL involvement in the conversations, and its negative effect on IC, appears to be a recurring theme in this study and will be discussed below.

## 8 Discussion

The results of this study showed that, when compared to the American group, the Japanese group sent backchannels far more frequently overall, which most notably included a greater percentage of SSBs and minimal responses. The American group, predictably, spoke a great deal more, posed a far greater number of questions and produced a greater percentage of extended responses. In addition, this study demonstrated a tendency of the Japanese participants to produce unconventional backchannels (such as continuer, understanding, agreement, and support and empathy type) in situations when they did not understand what their interlocutor was saying. Interpreting these results in this section, the researcher will examine how the above-mentioned differences influence IC, as well how the findings of this study compare to previous studies in the areas of frequency,

SSBs, conversational involvement, variability, discourse contexts, and function of backchannels.

## 8.1 Frequency

First, in dealing with frequency, previous studies conducted by Cutrone (2005) and White (1989) have used correlational analyses to test Lebra (1976) and Mizutani's (1982) hypothesis that NESs may take such frequent interjections as a sign of the listener's impatience and demand for a quick completion of the statement. This view was supported by the results of Cutrone (2005), who found that as JEFL backchannels increased, the more the British NESs felt they were being interrupted and the more they perceived their JEFL interlocutor to be impatient. The findings of this study, however, were similar to those reported by White (1989), who also found that more frequent JEFL backchannels were associated with positive ratings given by American NESs (observers in this study) on multiple items in the questionnaire. The overall positive ratings that the NES interlocutors gave to their JEFL counterparts (shown in Table 9) also suggest that the disparity in backchannel frequency did not greatly influence perceptions in a negative way. These contrary findings also highlight the fact that there may be considerable differences in how NESs of different varieties such as British English and American English perceive backchannels.

## 8.2 Simultaneous speech backchannels (SSBs)

The results presented in Table 8, showing that the JEFs' mean percentage of backchannels constituted by SSBs was noticeably greater than that of the NESs (1.7:1), were generally consistent with the results of Cutrone (2005) and Hayashi (1988) who found slightly greater ratios of 2.7:1 and 2.5:1, respectively. In light of Lebra (1976) and Mizutani's (1982) hypothesis that frequent interjections uttered by Japanese L2 English speakers are often ill-received by NESs, it was somewhat unexpected that more frequent (non-laughter) SSBs were associated with positive ratings on multiple items of the NES observers' questionnaire (see Appendix). Although the correlational analyses showed that SSBs might have actually had a positive effect on the intercultural conversations, the three NES interlocutors reported quite mixed feelings regarding this issue in the interviews. One NES was not even conscious of SSBs existing in the conversations, while the other two offered conflicting opinions. As Section 7.3 reported, one NES favored a great deal of listener feedback, while the other felt too much listener feedback could be intru-

sive and confounding, especially if he were in the midst of speaking. Considering the mixed findings here, and the lack of empirical data elsewhere, it is not prudent to draw any conclusions concerning Lebra (1976) and Mizutani's (1982) hypothesis at the present time. The great disparity in the three NESs' opinions regarding this issue seems to lend further support to the view that listenership, and how it is perceived, are highly individualistic and context-specific matters.

### 8.3 Conversational involvement

It is not surprising to observe the large number of backchannels produced by the JEFLs compared to how little they actually spoke in the intercultural conversations. Undoubtedly, it is clear that lack of involvement, and/or perceived lack of involvement, in the intercultural conversations played a major role in the NESs' perceptions of their JEFL interlocutors. While many of the JEFLs commented that they were generally more comfortable in a listener's role, all three NESs indicated that they would have liked their JEFL interlocutors to participate more in the conversations in terms of speaking more and asking more questions. The results in Section 6.1 show that the NESs spoke more than four times as many words as the JEFLs (264 and 62 words per conversation, respectively) and asked nearly twice as many questions (2.33 and 1.27 questions per conversation respectively). Although the NESs generally expected and accepted that they would have to carry the conversation in their NS-NNS exchanges, they also admitted that this onus detracted from their conversational satisfaction and enjoyment. These results were consistent with the findings of Cutrone (2005) and Sato (2008), who also found JEFLs' reticence as a source that would negatively influence cross-cultural perceptions and conversational satisfaction.

### 8.4 Variability

As the findings reported in Section 6.2 demonstrate, the NESs exhibited greater variability, overall, in the types, as well as the forms, of backchannels they employed as compared to the JEFL group. The JEFLs relied mainly on producing minimal responses and were not prone to elaboration, while the NESs uttered noticeably more extended responses (>27% difference in mean percentage of total backchannels). What this means in a conversation is likely to vary depending on the form and function of each backchannel form in its specific context; however, it is clear from the NESs' post-conversation interview responses in Section 7.3 that the JEFL's perceived lack of variability and exuberance was generally thought to

affect IC in a negative way, which is consistent with the observations made by Boxer (1993), Cutrone (2005), McCarthy (2003), and Stubbe (1998).

Concerning individual aspects of variability, the findings are able to shed further light on tendencies associated with verbal backchannel forms. While the NESs utilized a much greater percentage of isolated content words and multi-word phrases in their backchannels, minimal backchannels were still found to be the most common type of backchannel sent by both cultural groups. This was expected as various researchers such as McCarthy (2003) and Schegloff (1982) have alluded to a certain economy that seems to be built into oral communication. In other words, it is not at all unusual for speakers to utter no more than the bare minimum response. However, as McCarthy (2003) points out, speakers do not seem to economize when it comes to sociability. Generally speaking, varied responses coming at context-specific moments of the primary speaker's speech would perhaps go the farthest toward establishing active listenership and a positive effect in the conversation (Cutrone 2005; Stubbe 1998). On the other end of the spectrum, minimal and/or repetitive listener responses over an extended stretch of talk run the risk of being perceived as a sign of boredom or inattentiveness (McCarthy 2003), which is what some of NESs in this study perceived of their JEFLL interlocutors' listener responses.

Regarding the use of nonverbal backchannels, the JEFLLs and NESs were similar in most respects; however, the data revealed two marked differences. It was evident that the JEFLLs provided noticeably more head nods (1.6:1) and laughter (4.2:1) than the NESs. The former was expected as Maynard (1986, 1990, 1997) and Cutrone (2005) have shown that Japanese tend to nod much more frequently than NESs in IC. The latter was somewhat unexpected as Cutrone (2005) and Maynard (1997) found that the Japanese L2 English speakers in their studies articulated relatively similar amounts of laughter as the NESs. One possible explanation regarding the disparity in the amount of laughter in this study is that the JEFLLs used laughter as a way of reducing their anxiety and helping create a comfortable atmosphere in the conversation (as stated in Section 7.2).

## 8.5 Discourse contexts

Concerning discourse contexts favoring backchannels, there does appear to be some moderate variability between the two cultural groups; however, there was no evidence showing that the variability in discourse contexts negatively affected the conversations. The wide-ranging discourse contexts favoring the JEFLLs' backchannels seem largely to be a by-product of the great frequency by which they employ backchannels in conversations. That is, compared to the NESs, the JEFLLs'

backchannels were found to occur frequently and their discourse contexts favoring backchannels varied considerably; thus, it should be no surprise that many of the JEFs' backchannels were found at discourse contexts also prevalent in American English. Moreover, backchannels that are provided in discourse contexts other than the ones NESs commonly use are not necessarily considered inappropriate and are largely dependent on the context of the conversation and the function that the non-primary speaker desires to convey.

## 8.6 Function

One of the goals of this study was to examine the extent and reasons that intercultural miscommunication takes place in this context. In addition to conducting post-conversation playback interviews, the researcher included corresponding items in each group's questionnaires to delve deeper into this issue. The results pertaining to Item 2 in the questionnaires (outlined in Section 6.5) seem to suggest that the cross-cultural conversational interlocutors were not in-sync in some areas of the conversations and that some miscommunication may have occurred. While the NESs generally believed their JEFs' interlocutors to have represented themselves in a sincere and authentic manner in the conversations, the JEFs' scores acknowledged that they generally felt that they did not represent themselves in a sincere and authentic manner. This finding seems to lend support to the idea that Japanese L2 English speakers may sometimes feign understanding and/or agreement in order to keep conversations pleasant. This belief was further strengthened by several of the JEFs' admission in the playback interviews that they often employ continuer, understanding, agreement, and/or support and empathy type backchannels in situations when they did not understand (71%) and/or when they disagreed with what their interlocutor was saying (100%).

# 9 Conclusion

## 9.1 Summary

In answering the RQs of this study (see Section 4), this study provides the most comprehensive evidence to date that JEFs' listenership differs to that of NESs in many respects (RQ 1), and, by examining the effect these differences had on IC, has also demonstrated that these differences sometimes lead to miscommunication, negative perceptions, and stereotyping (RQ 2). Considering the importance

of listenership in IC, and the fact that it is currently not included in the EFL curriculum in Japan (Capper 2000; Cutrone 2010; Okushi 1990), it would seem that the next logical step in the research is to ascertain how this key aspect of pragmatic competence can be better acquired by JEFLL learners.

## 9.2 Pedagogical implications

Besides providing justification for the teaching of listenership in JEFLL classes, the findings of this study can be the first steps toward establishing a pedagogical framework in this area. That is, the conversational behavior of proficient speakers of English (in this case, American NESs), as documented in the transcribed conversations, can supplement models for adequate listenership that are currently being developed (Cutrone 2010). Although the use of American NES norms as a model for backchannel behavior is presently justified in this paper (as there are no other concrete and comprehensive options currently available in which backchannels are concerned), it would be a positive development to see researchers one day piece together a more thorough and representative description of English backchannel behavior as it exists in the international community. By doing this, a broader set of goals and norms could be presented in EFL/ESL contexts around the world. Nonetheless, it is important to point out that the instructional goals for EFL/ESL backchannel behavior will probably always have much more to do with the recognized tendencies and idiosyncrasies of a particular group of learners than following the prescriptive norms of any one NES variety and/or international standard. Generally speaking, the goals for JEFLLs should involve sending minimal backchannels less frequently (especially while one's interlocutor is speaking), with greater variability (but at context-appropriate moments), while asking questions and taking the primary speakership in the conversation more often, and initiating conversational repair strategies when they do not understand and/or disagree rather than feign understanding and agreement. Relating to this final point, this study has touched upon the sensitive issue of teaching culture in the classroom. That is, as many of the JEFLLs have expressed, the basis for their listening behavior may involve the different cultural values and codes of conduct they are accustomed to following (Cutrone 2011).

Although differing functional objectives of listenership may certainly contribute to miscommunication in IC, EFL teachers would be well advised to not push or force their learners to use and/or avoid using backchannels in ways where they may feel uncomfortable. Many learners are more than willing to embrace the target culture in their efforts to learn English, and for those who hold back initially, it is likely that comfort levels will increase gradually over time as

learners become more acclimated to the target culture. Teachers can facilitate this process in two ways: (1) by initiating activities to raise learners' consciousness regarding the observed differences of listenership across cultures and (2) by helping learners deconstruct the potential reasons underpinning such conversational behavior. Ultimately however, the researcher believes that instruction in this area should go beyond the raising-awareness phase to subsequently include practical activities that provide opportunities for output and corrective feedback.

## References

- Blanche, Patrick. 1987. The case for a pedagogy of pragmatics in foreign or second language teaching. *Regional Language Center (RELC) Journal* 18(1). 46–71.
- Brunner, Lawrence J. 1979. Smiles can be backchannels. *Journal of Personality and Social Psychology* 37(1). 728–734.
- Clancy, Patricia M., Sandra A. Thompson, Ryoko Suzuki, & Hongyin Tao. 1996. *The conversational use of reactive tokens in English, Japanese, and Mandarin*. *Journal of Pragmatics* 26(1). 355–387.
- Crawford, William. 2003. Back-channel transfer in cross-cultural second language conversation: Do native Japanese transfer more than native English speakers? *Journal of Kitakyushu National College of Technology* 36(1). 155–163.
- Cutrone, Pino. 2005. A case study examining backchannels in conversations between Japanese-British dyads. *Multilingua: Journal of Cross-Cultural and Interlanguage Communication* 24(1). 237–274.
- Cutrone, Pino. 2010. The backchannel norms of native English speakers: A target for Japanese L2 English learners. *University of Reading Language Studies Working Papers* 2(1). 28–37.
- Cutrone, Pino. 2011. *Towards the learning of listener responses in the Japanese EFL context*. PhD Dissertation, University of Reading, Reading.
- Cutrone, Pino. Pending. Explicit vs. implicit instruction: Investigating pragmatic development in the Japanese EFL classroom. *Asian EFL Journal*.
- Dittman, Allen T. & Lynn G. Llewellyn. 1968. Relationship between vocalization and head nods as listener responses. *Journal of Personality and Social Psychology* 9(1). 79–84.
- Duncan, Starkey Jr. 1974. On the structure of speaker-auditor interaction during speaking turns. *Language in Society* 2(1). 161–180.
- Duncan, Starkey Jr. & Donald Fiske. 1977. *Face to face interaction: Research, methods, and theory*. Hillsdale: Erlbaum.
- Ellis, Rod. 1991. Communicative competence and the Japanese learner. *The Japan Association for Language Teachers (JALT) Journal* 13(1). 103–127.
- Farooq, Mohammad U. 2005. A model for motivating Japanese EFL learners through real-life questioning strategies. *The Journal of Nagoya Gakugei University* 1(1). 27–42.
- Fries, Charles. 1952. *The structure of English*. New York: Harcourt Brace.
- Fujimoto, Donna T. 2007. Listener responses in interaction: A case for abandoning the term, backchannel. *Journal of Osaka Jogakuin College* 37(1). 35–54.
- Hall, Edward T. 1974. *Handbook for proxemic research*. Washington: Society for the Anthropology of Visual Communication.

- Hattori, Takahiko. 1987. A study of nonverbal intercultural communication between Japanese and Americans focusing on the use of the eyes. *The Japan Association for Language Teachers (JALT) Journal* 8(1). 109–118.
- Hayashi, Reiko. 1988. Simultaneous talk: From the perspective of floor management of English and Japanese speakers. *World Englishes* 7(3). 269–288.
- Hecht, Michael L. 1978. The conceptualization and measurement of interpersonal communication satisfaction. *Human Communication Research* 4(3). 253–264.
- Helgesen, Mark. 1987. False beginners: Activating language for accuracy and fluency. *The Language Teacher* 11(14). 23–29.
- Hill, Thomas. 1990. Sociolinguistic aspects of communicative competence and the Japanese learner. *Dokkyo University Studies in English* 36(1). 69–104.
- Hughes, Henry J. 1999. Cultivating the walled garden: English in Japan. *English Studies* 80(6). 556–568.
- Ike, Saya. 2010. Backchannel: A feature of Japanese English. In Alan M. Stoke (ed.), *The Japan Association for Language Teachers (JALT) 2009 Conference Proceedings*. Tokyo: JALT. 205–215.
- Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26(1). 22–63.
- Kendon, Adam. 1977. Some functions of gaze-direction in two-person conversation. In Adam Kendon (ed.), *Studies in the Behavior of Social Interaction*, 13–51. Bloomington: Indiana University Press.
- Lebra, Takie S. 1976. *Japanese patterns of behavior*. Honolulu: The University of Hawaii Press.
- LoCastro, Virgiana. 1987. Aizuchi: A Japanese conversational routine. In Larry Smith (ed.), *Discourse across cultures: Strategies in world Englishes*, 101–113. London: Prentice Hall.
- Markel, Norman N. 1975. Coverbal behavior associated with conversation turns. In Adam Kendon, Richard M. Harris, & Mary Ritchie Key (eds.), *Organization of behavior in face-to-face interaction*, 189–197. The Hague: Mouton.
- Matsumoto, Kaoru. 1994. English instruction problems in Japanese schools and higher education. *Journal of Asian Pacific Communication* 5(4). 209–214.
- Maynard, Senko K. 1986. On back-channel behavior in Japanese and English casual conversation. *Linguistics* 24(6). 1079–1108.
- Maynard, Senko K. 1987. Interactional functions of a nonverbal sign: Head movement in Japanese dyadic conversations. *Journal of Pragmatics* 11(1). 589–606.
- Maynard, Senko K. 1989. *Japanese conversation: Self-contextualization through structure and interactional management, Volume 35*. Norwood: Ablex.
- Maynard, Senko K. 1990. Conversation management in contrast: Listener responses in Japanese and American English. *Journal of Pragmatics* 14(1). 397–412.
- Maynard, Senko K. 1997. Analyzing interactional management in native/non-native English conversation: A case of listener response. *International Review of Applied Linguistics in Language Teaching (IRAL)* 35(1). 37–60.
- McCarthy, Michael. 2002. Good listenership made plain: British and American non-minimal response tokens in everyday conversation. In Randi Reppen, Susan M. Fitzmaurice, & Douglas Biber (eds.), *Using corpora to explore linguistic variation*, 49–72. Philadelphia: John Benjamins.
- McCarthy, Michael. 2003. Talking back: “Small” interactional response tokens in everyday conversation. *Research on Language and Social Interaction* 36(1). 33–63.



- McCarthy, Michael & Anne O’Keeffe. 2004. Research in the teaching of speaking. *Annual Review of Applied Linguistics* 24(1). 26–43.
- McCroskey, James C. 1992. Reliability and validity of willingness to communicate scale. *Communication Quarterly* 40(1). 16–25.
- Mizutani, Nobuko. 1982. The listener’s responses in Japanese conversation. *Sociolinguistics News-letter* 13(1). 33–38.
- Okushi, Yoshiko. 1990. Misunderstood efforts and missed opportunities: An examination of EFL in Japan. *Penn Working Papers in Educational Linguistics* 6(2). 65–75.
- Oreström, Bengt. 1983. *Turn-Taking in English Conversation*. Lund: Lund University Press.
- Reesor, Mathew. 2002. The bear and the honeycomb: A history of Japanese English language policy. *Nagoya University of Commerce and Business Journal of Language, Culture and Communication (NUCB JLCC)* 4(1). 41–52.
- Roger, Alun. 2008. Teaching the speaking skill to Japanese students part 1: Construct & practice. *The Journal of Kanda University of International Studies* 20(1). 1–26.
- Ross, Steven. 1998. Divergent frame interpretations in oral proficiency interview interaction. In Richard Young & Agnes Weiyun He (eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency*, 333–353. Philadelphia: John Benjamins.
- Sacks, Harvey E., Emanuel A. Schegloff, & Gail Jefferson. 1974. A simple systematics in the organization of turn-taking for conversation. *Language* 50(1). 696–735.
- Sato, Yoko. 2008. *Oral communication problems and strategies of Japanese university EFL learners*. PhD Dissertation, University of Reading, Reading.
- Schegloff, Emanuel A. 1982. Discourse as an interactional achievement: Some uses of “UH-HUH” and other things that come between sentences. In Deborah Tannen (ed.), *Georgetown University roundtable on language and linguistics, analyzing discourse: Text and talk*, 71–93. Washington: Georgetown University Press.
- Schenkein, James N. 1972. Towards the analysis of natural conversation in the sense of heheh. *Semiotica* 6(1). 344–377.
- Stubbe, Maria. 1998. Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics* 29(3). 257–289.
- Takanashi, Yoshiro. 2004. TEFL and communication styles in Japanese culture. *Language, Culture and Curriculum* 17(1). 1–14.
- Tottie, Gunnel. 1991. Conversational style in British and American English: The case of backchannels. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 254–271. London: Longman.
- Ward, Nigel & Tsukahara, Wataru. 2000. Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics* 32(8). 1177–1207.
- White, Shieda. 1989. Backchannels across cultures: A study of Americans and Japanese. *Language in Society* 18(1). 59–76.
- Yano, Yasukata. 2001. World Englishes in 2000 and beyond. *World Englishes* 20(2). 119–131.
- Yashima, Tomoko. 2002. Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal* 86(1). 54–66.
- Yngve, Victor. 1970. On getting a word in edgewise. *Chicago Linguistic Society* 6(1). 567–578.

## Appendix

### Spearman rho correlational analyses and explanations

Key for Terms: BCs = Backchannels, SSBs = Simultaneous Speech Backchannels (not including laughter), EXBCs = Extended Backchannels, MINBCs = Minimal Backchannels, WORDS = Number of Words

Correlations between the JEFsLs' frequencies of each conversational feature (shown on the left) and the NES observers' ratings on each item of the conversational satisfaction questionnaire (shown on the right) are shown as follows: \* =  $p < .05$ , \*\* =  $p < .01$

Spearman's rho		ITEM01	ITEM02	ITEM03	ITEM04	ITEM05	ITEM06
BC	Correlation Coefficient	-.326	-.398(*)	-.260	-.283	-.415(*)	-.530(**)
	Sig. (2-tailed)	.078	.029	.166	.130	.023	.003
	N	30	30	30	30	30	30
SSB	Correlation Coefficient	-.615(**)	-.575(**)	-.439(*)	-.678(**)	-.554(**)	-.682(**)
	Sig. (2-tailed)	.001	.001	.015	.001	.001	.001
	N	30	30	30	30	30	30
EXBC	Correlation Coefficient	-.498(**)	-.468(**)	-.340	-.523(**)	-.364(*)	-.466(**)
	Sig. (2-tailed)	.005	.009	.066	.003	.048	.009
	N	30	30	30	30	30	30
MINBC	Correlation Coefficient	-.206	-.284	-.178	-.153	-.315	-.342
	Sig. (2-tailed)	.275	.129	.346	.419	.090	.064
	N	30	30	30	30	30	30
WORD	Correlation Coefficient	-.366(*)	-.383(*)	-.233	-.345	-.179	-.167
	Sig. (2-tailed)	.047	.037	.216	.062	.344	.378
	N	30	30	30	30	30	30

Spearman's rho		ITEM07	ITEM08	ITEM09	ITEM10	ITEM11	ITEM12
BC	Correlation Coefficient	-.253	.352	.328	-.220	-.277	.341
	Sig. (2-tailed)	.177	.056	.077	.243	.139	.065
	N	30	30	30	30	30	30
SSB	Correlation Coefficient	-.489(**)	.375(*)	.250	-.343	-.498(**)	.449(*)
	Sig. (2-tailed)	.006	.041	.183	.064	.005	.013
	N	30	30	30	30	30	30
EXBC	Correlation Coefficient	-.341	.240	.324	-.334	-.605(**)	.325
	Sig. (2-tailed)	.065	.202	.081	.071	.001	.080
	N	30	30	30	30	30	30
MINBC	Correlation Coefficient	-.185	.307	.215	-.169	-.081	.311
	Sig. (2-tailed)	.329	.099	.255	.373	.669	.095
	N	30	30	30	30	30	30
WORD	Correlation Coefficient	-.477(**)	-.068	.283	-.136	-.278	.146
	Sig. (2-tailed)	.008	.720	.130	.472	.137	.441
	N	30	30	30	30	30	30

Spearman's rho		ITEM13	ITEM14	ITEM15	ITEM16	ITEM17
BC	Correlation Coefficient	-.290	-.265	.185	-.173	.306
	Sig. (2-tailed)	.121	.157	.329	.360	.100
	N	30	30	30	30	30
SSB	Correlation Coefficient	-.607(**)	-.250	.400(*)	-.460(*)	.526(**)
	Sig. (2-tailed)	.001	.183	.028	.011	.003
	N	30	30	30	30	30
EXBC	Correlation Coefficient	-.493(**)	-.169	.468(**)	-.416(*)	.561(**)
	Sig. (2-tailed)	.006	.372	.009	.022	.001
	N	30	30	30	30	30
MINBC	Correlation Coefficient	-.098	-.174	.041	-.050	.187
	Sig. (2-tailed)	.607	.357	.829	.792	.323
	N	30	30	30	30	30
WORD	Correlation Coefficient	-.343	-.153	.591(**)	-.473(**)	.268
	Sig. (2-tailed)	.064	.419	.001	.008	.152
	N	30	30	30	30	30

Key explaining significant findings above:

### BCs

- Item 2: As the number of JEFL backchannels increased, the more the NES observers believed the JEFLs to be showing understanding ( $p < .029$ ).
- Item 5: As the number of JEFL backchannels increased, the more the NES observers felt the conversation went smoothly ( $p < .023$ ).
- Item 6: As the number of JEFL backchannels increased, the more the NES observers believed them to be encouraging their interlocutor to continue speaking ( $p < .003$ ).

### SSBCs

- Item 1: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the JEFLs were letting their partner know they were communicating effectively ( $p < .001$ ).
- Item 2: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the JEFLs were showing that they understood their partner ( $p < .001$ ).
- Item 3: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the JEFLs were listening attentively to what their partner said ( $p < .015$ ).
- Item 4: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the JEFLs expressed a lot of interest in what their partner had to say ( $p < .001$ ).
- Item 5: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the conversation went smoothly ( $p < .001$ ).
- Item 6: As the number of JEFL non-laughter SSBs increased, the more the NES observers believed the JEFLs to be encouraging their interlocutor to continue speaking ( $p < .001$ ).
- Item 7: As the number of JEFL non-laughter SSBs increased, the more the NES observers believed the JEFLs to be expressing themselves in an authentic and sincere manner ( $p < .006$ ).
- Item 8: As the number of JEFL non-laughter SSBs increased, the less the NES observers felt the JEFLs to be impatient ( $p < .041$ ).
- Item 11: As the number of JEFL non-laughter SSBs increased, the more the NES observers felt the JEFLs to be warm and friendly ( $p < .005$ ).
- Item 12: As the number of JEFL non-laughter SSBs increased, the less the NES observers believed the JEFLs to be impolite ( $p < .013$ ).
- Item 13: As the number of JEFL non-laughter SSBs increased, the more the NES observers believed the JEFLs to be interested and concerned ( $p < .001$ ).

- Item 15: As the number of JEFL non-laughter SSBs increased, the less the NES observers believed the JEFLs to want to avoid speaking ( $p < .028$ ).
- Item 16: As the number of JEFL non-laughter SSBs increased, the more the NES observers believed the JEFLs were able to express that they did not understand ( $p < .011$ ).
- Item 17: As the number of JEFL non-laughter SSBs increased, the less the NES observers believed the JEFLs' listening behavior to be inadequate ( $p < .003$ ).

### EXBCs

- Item 1: As the number of JEFL extended responses increased, the more the NES observers felt the JEFLs were letting their partner know they were communicating effectively ( $p < .005$ ).
- Item 2: As the number of JEFL extended responses increased, the more the NES observers felt the JEFLs were showing that they understood their partner ( $p < .009$ ).
- Item 4: As the number of JEFL extended responses increased, the more the NES observers felt the JEFLs expressed a lot of interest in what their partner had to say ( $p < .003$ ).
- Item 5: As the number of JEFL extended responses increased, the more the NES observers felt the conversation went smoothly ( $p < .048$ ).
- Item 6: As the number of JEFL extended responses increased, the more the NES observers believed the JEFLs to be encouraging their interlocutor to continue speaking ( $p < .009$ ).
- Item 11: As the number of JEFL extended responses increased, the more the NES observers felt the JEFLs to be warm and friendly ( $p < .001$ ).
- Item 13: As the number of JEFL extended responses increased, the more the NES observers believed the JEFLs to be interested and concerned ( $p < .006$ ).
- Item 15: As the number of JEFL extended responses increased, the less the NES observers believed the JEFLs to want to avoid speaking ( $p < .009$ ).
- Item 16: As the number of JEFL extended responses increased, the more the NES observers believed the JEFLs were able to express that they did not understand ( $p < .022$ ).
- Item 17: As the number of JEFL extended responses increased, the less the NES observers believed the JEFLs' listening behavior to be inadequate ( $p < .001$ ).

### MINBCs

No significant findings were observed.

WORDS

- Item 1: As the number of JEFL words increased, the more the NES observers felt the JEFLs were letting their partner know they were communicating effectively ( $p < .047$ ).
- Item 2: As the number of JEFL words increased, the more the NES observers felt the JEFLs were showing that they understood their partner ( $p < .037$ ).
- Item 7: As the number of JEFL words increased, the more the NES observers believed the JEFLs to be expressing themselves in an authentic and sincere manner ( $p < .008$ ).
- Item 15: As the number of JEFL words increased, the less the NES observers believed the JEFLs to want to avoid speaking ( $p < .001$ ).
- Item 16: As the number of JEFL words increased, the more the NES observers believed the JEFLs were able to express that they did not understand ( $p < .008$ ).