Backchannels, Appraisal Psychology, and Second Language Fluency

JAMES P. WOLF*

Abstract

The growth of global tourism necessitates a better understanding of factors that affect intercultural communication. One factor concerns how people from different sociocultural and linguistic environments use and react to backchannels. Backchannels are verbal and nonverbal messages sent by an interlocutor playing a listener’s role in conversation. Verbal backchannels, for example, consist of ‘uh-huh’, ‘mm-hm’, or ‘yeah’ while nonverbal messages involve head nods. Sociocultural and linguistic groups have been found to differ with respect to the frequency, placement, functions and types of backchannels used. To highlight these differences, the backchannel behavior of Japanese and American English speakers is contrasted in this paper. Appraisal psychology is discussed to explain how reactions to backchannels may vary across sociocultural and linguistic groups. Then the backchannel output hypothesis is proposed to predict how the fluency of L2 learners of English may be affected by their reactions to English backchannels during classroom oral tasks.

Key words: Backchannels, Appraisal psychology, Second language fluency, Language teaching, Intercultural communication

1. Introduction

With the continued rise in global tourism, there is an even greater need to understand factors that may contribute to or inhibit successful intercultural communication. One such factor concerns how people from different sociocultural and linguistic backgrounds use and react to backchannels. Backchannels are verbal

*Faculty of Regional Development Studies, Toyo University, Japan
and nonverbal messages sent by an interlocutor playing a listener’s role in conversation. Verbal messages may consist of, for example, ‘uh-huh’, ‘mm-hm’, and ‘yeah’, while head nods and smiles are examples of nonverbal messages. In the literature, such messages for expressing listenership have received numerous labels such as ‘signals of attention’ (Fries, 1952), ‘accompaniment signals’ (Kendon, 1967), ‘listener responses’ (Dittmann and Llewellyn, 1968) and ‘backchannels’ (Yngve, 1970). In this paper, I shall refer to them interchangeably as backchannels or backchannel cues.

Backchannels constitute a universal human behavior; however, specific backchannel behaviors are particular to language and culture (Heinz, 2003). Much research has shown that sociocultural and linguistic groups differ with respect to the (1) frequency, (2) placement, (3) functions and (4) types of backchannel cues used (e.g. Clancy et al., 1996; Maynard, 1990, 1997; White, 1989). These differences can lead to misunderstandings during intercultural communication. For example, Cutrone (2005) investigated the effects of backchannels on intercultural communication in eight dyadic English conversations between British and Japanese participants and discovered that the British participants had negative perceptions of their Japanese counterparts. In particular, the British participants complained that the Japanese seemed impatient because they used backchannels too frequently. They also felt that the Japanese overused backchannels to avoid speaking due to worries about their English ability and sent backchannels even when disagreeing or not understanding.

In this paper, rather than focusing on how different sociocultural and linguistic groups use backchannels, I will instead examine how they react to backchannel cues. More specifically, my discussion will focus on second language (L2) learners of English and how they may react differently to English backchannel cues during classroom oral tasks. The remainder of this paper will be divided into the following sections. Firstly, I will contrast the backchannel behavior of Japanese and American English speakers in order to demonstrate how sociocultural and linguistic groups can differ with respect to this aspect of communication. Secondly, I will discuss Appraisal Psychology which can explain how L2 learners of English may react differently to English backchannel cues in socioculturally determined ways. Thirdly, I will propose the ‘backchannel output hypothesis’ which suggests that L2 learners of English from different first language (L1) and sociocultural backgrounds will show different levels of fluency during oral tasks in response to
English backchannel cues. Fourthly, I will define second language (L2) fluency and discuss research in this area. Lastly, I will conclude by summarizing this paper and offering suggestions for future research involving the backchannel output hypothesis.

2. Contrasting Japanese and American English Speakers’ Backchannel Behavior

How does backchannel use differ between Japanese and American English speakers? To answer this question, I will examine the (1) frequency, (2) placement, (3) functions, and (4) types of backchannels used between these two cultural and linguistic groups.

Several researchers have pointed out that Japanese speakers use backchannels much more frequently compared to American English speakers (e.g. Clancy et al., 1996; LoCastro 1987, 1999; Maynard, 1990, 1997; White, 1989). The Japanese participants in White's (1989) study uttered one backchannel for every 14 words while speaking English with Americans who uttered one for every 37 words. White noted that this 3:1 ratio remains when Japanese and Americans speak with fellow native speakers in their respective languages. Why do Japanese speakers use backchannel cues so extensively? Several researchers have mentioned Japanese cultural influences, connecting their relatively frequent use to social values such as consideration of others, cooperation, and the desire for harmonious communication (e.g. LoCastro, 1987, 1999; Maynard 1997; White, 1989). Others have taken a linguistic perspective, suggesting that discourse structures inherent in Japanese naturally elicit backchannel responses (Maynard, 1997; Miller, 1988; White, 1989). In contrast, LoCastro (1999) proposed a functional perspective, claiming that the language code has evolved due to cognitive and socio-cultural pressures to allow more frequent backchannel cues.

Secondly, Japanese and American English speakers also differ concerning where they place backchannels. American English speakers consistently tend to produce backchannels at points of grammatical completion and thus do not usually overlap their interlocutors' talk with them (Clancy et al., 1996; LoCastro, 1987; Maynard, 1990, 1997). Conversely, the discourse contexts of Japanese speakers' backchannels are varied. Maynard (1990) reported that her Japanese participants' backchannels occurred at grammatical junctures, during unfilled pauses, and were also prompted by cues from the primary speaker such as sentence final particles and
head nods. Furthermore, Japanese speakers are more likely to send backchannels
during the primary speaker’s turn thus producing simultaneous talk (e.g. Clancy et al., 1996; LoCastro, 1987; Maynard, 1990, 1997).

Thirdly, functional differences also exist between Japanese and American English speakers in their backchannel use. These differences, however, are not as
great as those regarding the frequency and placement of backchannel cues already discussed. In fact, based on her analysis, Maynard (1997, p.46) reported that both
Japanese and American English speakers use backchannels for the same purposes:
(1) continuer (Schegloff, 1982); (2) display of understanding of content; (3)
support toward the speaker’s judgment; (4) agreement; (5) strong emotional
response; (6) minor addition, correction, or request for information. Although
these functions can overlap and are not mutually exclusive, Maynard observed that
Japanese favor the display of understanding of content function to express moral
support for the speaker, while Americans favor the ‘continuer’ function (i.e., to send
a backchannel to pass up an opportunity for a speaking turn at a grammatical
juncture).

Lastly, somewhat greater similarity also exists concerning the types of
backchannels used. Maynard’s (1990) contrastive analysis of 40 (20 American and
20 Japanese) dyadic conversations found brief utterances (i.e., ‘uh-huh’, ‘mm-hm’,
etc.) the most frequent type used by both Japanese and American English speakers
(70.49% and 50.23%, respectively) and head movement (i.e., head nods) the second
most frequent (18.83% and 35.05%, respectively). Regarding brief utterances,
however, LoCastro (1987, 1999) stated that there appears to be a greater variety of
these in Japanese than in American English. Overall, Maynard (1990) declared the
biggest differences between Japanese and American English backchannels to be in
their frequency and discourse contexts, as discussed earlier.

3. Appraisal Psychology

Appraisal theorists contend that emotions are generated and distinguished
through a process of stimulus appraisal in which an individual subjectively
evaluates or appraises agents, objects, and situations in its environment (Leventhal
Several models of the psychological dimensions along which these appraisals are
made have been developed; however, Scherer’s (1984) model is considered most
comprehensive (Schumann & Wood, 2004). Scherer (1984) views stimuli as being evaluated according to their (1) novelty, (2) intrinsic pleasantness, and their compatibility (3) with one’s goals/needs, (4) with one’s coping potential, and (5) with one’s self and social image.

Leventhal and Scherer (1987) suggested a levels of processing approach in which Scherer’s (1984) five psychological dimensions for stimulus appraisal can be processed at a sensorimotor, schematic, or conceptual level as shown in Table 1. According to Scherer (1987), appraisals at the sensorimotor level are made via hard-wired detection features and reflex systems. Let’s look at an example of an appraisal at the sensorimotor level along the psychological dimension of novelty. As seen in Table 1, the novelty appraisal can involve suddenness. If someone were to stealthily sneak up behind you on the street and suddenly scream, you would undoubtedly be startled and your appraisal of the stimuli (i.e. the scream) would generate a feeling of fear. In essence, your psychological evaluation of the external stimuli would produce a physical emotion.

Similarly, as shown in Table 1, appraisals along the dimension of pleasantness at the sensorimotor level are made on the basis of innate preferences and aversions labeled by Schumann (1997) as homeostatic value. Schumann explains that such preferences promote an organism’s survival and thriving in the world since homeostatic systems control respiration, heart rate, body temperature, satiation, eliminatory functions, and sexual drive. As an example of an appraisal of pleasantness at the sensorimotor level, Scherer (1987) notes that many animals as well as infants and humans have a hard-wired preference for sweet and an aversion for bitter tastes.

<table>
<thead>
<tr>
<th>Table 1 Processing Levels for Stimulus Evaluation Checks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual Level</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Expectations cause/effect probability estimates</td>
</tr>
<tr>
<td>Schematic Level</td>
</tr>
<tr>
<td>Sensorimotor Level</td>
</tr>
</tbody>
</table>

Note. Reproduced from Leventhal & Scherer, 1987, p.17
While some preferences and aversions are innate, others are learned as shown along the psychological dimension of pleasantness at the schematic level in Table 1. Schumann (1997) describes these learned preferences and aversions as somatic value which is acquired in the lifetime of the individual through experience, socialization, enculturation, and education.

These preferences and aversions are composed of schemata which can be conceptualized as abstract representations of learned responses to specific stimulus patterns (Scherer, 1987). It is along the pleasantness dimension at the schematic level where L2 learners of English from various L1 and sociocultural backgrounds can be expected to differ with respect to their preferences for receiving English backchannel cues during their talk. It should be mentioned that the appraisals at both the sensorimotor and schematic levels are considered by Appraisal theorists to occur in a rapid fashion outside of conscious awareness. At the conceptual level, however, appraisals become increasingly complex as seen in Table 1. In contrast to the sensorimotor and schematic levels, appraisals here are mostly considered to involve conscious awareness.

4. The Backchannel Output Hypothesis

Tao and Thompson (1991) studied two native Mandarin speakers for whom American English had become their dominant language after having lived in the United States for many years. Tao and Thompson found that they pragmatically transferred their L2 American English backchannel behavior to their L1 Mandarin conversations. Heinz (2003) later studied native German speakers and also found that these individuals transferred their L2 American English backchannel behavior to their L1 German conversations after developing a high level of competence in American English. These studies therefore show that L2 learners of English eventually develop implicit knowledge of English backchannel cues upon reaching advanced levels of oral proficiency in the target language (TL).

In contrast, non-advanced English learners tend to transfer their L1 backchannel behavior to their L2 conversations. For example, both White (1989) and Maynard (1997) found the Japanese L2 learners of English in their respective studies to transfer their L1 backchannel behavior to their L2 English conversations. Maynard concluded that Japanese learners of English use backchannels in intercultural discourse in the same manner as within their own cultural context.
Rather than being restricted to Japanese learners of English, in my view non-advanced learners of English in general, due to their lack of implicit knowledge of English backchannel behavior, will generally have an implicit (i.e. unconscious) preference for producing and receiving backchannels in a manner consistent with the backchannel conventions of their respective L1 and socio-cultural environments.

Intimately associated with an individual’s implicit preference for how an interlocutor should respond to their talk with verbal and nonverbal backchannel cues is the value that is attached to such listener behaviors. It seems likely that non-advanced English learners from different socio-cultural and linguistic backgrounds will react to English backchannel cues in ways that reflect different appraisals of their value. As mentioned earlier, it is suggested that these appraisals will occur unconsciously at the schematic level along the dimension of pleasantness (see Table 1). As noted, these appraisals are made on the basis of learned preferences and aversions acquired in the lifetime of the individual through experience, socialization, enculturation, and education. Since specific backchannel behaviors are particular to culture and language, individuals from different sociocultural and linguistic environments will tend to have different sets of preferences for producing and receiving backchannels.

As discussed earlier, Schumann (1997) refers to learned preferences and aversions as somatic value. Such preferences influence an individual’s appraisal of stimuli along the psychological dimensions identified by Scherer (1984). From this perspective, backchannels might be considered emotional stimuli that L2 speakers evaluate via their somatic value systems and that may, depending on their preferences, facilitate their fluency. Thus, I would like to propose the backchannel output hypothesis, which suggests that non-advanced learners of English from various sociocultural and L1 contexts will tend to evaluate verbal and nonverbal English backchannel behavior differently based on their implicit preferences for receiving backchannels, and these evaluations will determine whether or not English backchannel cues facilitate their fluency during oral tasks.

5. Second Language (L2) Fluency

5.1.1. Defining L2 fluency

Defining L2 fluency has been a difficult task for researchers due to conflicting meanings associated with the term. In an effort to disambiguate the meaning of
fluency in foreign language learning, Lennon (1990) distinguished between a “broad” sense and a “narrow” sense of fluency. According to the broad sense, he noted that fluency serves as a cover term for oral proficiency, representing “the highest point on a scale that measures spoken command of a foreign language” (p. 389). On the other hand, Lennon observed that, in its narrow sense, fluency in EFL pertains to one, isolatable component of oral proficiency describing learners who are “fluent but grammatically inaccurate” or “fluent but lacks a wide and varied vocabulary”. Furthermore, Lennon (1990) mentioned that the overall emphasis on fluency in the narrow sense seems to be on native-like rapidity.

A focus on native-like rapidity dovetails well with Schmidt’s (1992) conception of L2 fluency. Schmidt defines second language fluency as a “procedural skill”, and states that an extreme example of such a fluent speaker “would be the speaker of a pidginized interlanguage, largely unanalyzed and agrammatic with respect to the target language, who speaks that variety in a fluid rather than a halting manner” (p.358). Researchers sharing such a view of L2 fluency operationalize and measure it in terms of various temporal and dysfluency markers such as speech rate, the length of pauses, the length of speech runs between pauses, and the number of repetitions and reformulations (e.g. Derwing et al., 2004; Freed, 1995; Kormos and Dénes, 2004; Lennon, 1990; Riggenbach, 1991; Towell et al., 1996).

5.1.2. Investigations of factors influencing L2 fluency during oral tasks

In recent years, several researchers have attempted to elucidate factors that influence L2 speakers’ fluency during oral tasks. Some have investigated the effects of task characteristics themselves, such as task types (Skehan and Foster, 1999; Derwing et al., 2004). Others have examined how task conditions, such as preplanning and online planning, affect fluency (e.g. Crookes, 1989; Foster and Skehan, 1996; Mehnert, 1998; Ortega, 1999; Wigglesworth, 1997; Yuan and Ellis, 2003). Aside from tasks, some researchers have investigated the effects of L2 speaker performance and speech variables, such as self-monitoring (Kormos, 1999) and intonation (Wennerstrom, 2000), while still others have studied how time spent in an L2 learning context may affect fluency (Freed, 1995; Freed et al., 2004; Lennon, 1990). In sum, fluency has been investigated with respect to three different variables: (1) task characteristics, (2) L2 speaker characteristics, and (3) learning context.

At least one important variable that may affect L2 speakers’ fluency during oral
tasks has eluded consideration; that is, the listener behaviors of their interlocutors. When L2 speakers perform oral tasks, teachers or testers are often present and respond to their production with a variety of verbal and nonverbal backchannel messages. In spite of this research gap in the L2 field, experimental research has been done in the psychotherapy field in an English L1 context to assess how interviewer backchannel cues affect interviewees’ verbal productivity. These studies aimed to show that patients could be induced to speak more during psychotherapy by responding to them with backchannel cues. Some studies (Kanfer and McBrearty, 1962; Matarazzo et al., 1964a, 1964b) found verbal and nonverbal backchannels to facilitate verbal productivity while others did not (Siegman, 1976). These results cannot, however, be used to infer how L2 speakers might react to backchannel cues since they involved L1 speakers in interviews rather than L2 speakers in tasks. Understanding how backchannels might affect L2 learners’ fluency is especially important for teachers who interact and negotiate meaning with them in the target language, as well as for language testers who evaluate their oral production.

5.1.3. An investigation of the effects of backchannels on fluency in L2 oral task production

In an effort to fill this research gap in the L2 field, Wolf (forthcoming) investigated the effect of different conditions of listener backchannels on the fluency of L2 speakers. Fluency was defined as an automatic procedural skill (Schmidt, 1992). Participants were 14 non-advanced Japanese learners of English who each performed three oral narrative tasks in three different backchannel conditions: (a) verbal/nonverbal (V/NV), (b) nonverbal-only (NV), and (c) no backchannels (NB). In the V/NV condition, both verbal and nonverbal backchannels were given in response to the Japanese participants’ oral task production. The verbal backchannels included ‘mm-hm’ and ‘uh-huh’ and the nonverbal backchannels involved head nodding. In the NV condition, only nonverbal backchannels (head nodding) were given. In the NB condition, neither verbal nor nonverbal backchannels were given. The Japanese participants’ fluency was measured via five temporal measures. Wolf hypothesized that the Japanese participants would be least fluent in the no-backchannel (NB) condition, more fluent in the nonverbal (NV) condition and most fluent in the verbal/nonverbal (V/NV) condition.

A series of one-way within-subjects ANOVAs were performed on all five
fluency measures followed by post-hoc t-tests. The alpha for achieving statistical
significance was set at .05 for the ANOVAs and at .017 for the post-hoc tests. The
Appendix shows that the hypotheses for the three backchannel conditions were all
supported by the fluency results for: (1) Rate A (the mean number of syllables
spoken per minute); (2) Rate B (same as Rate A, but excluding all syllables, words,
phrases that were repeated, reformulated, or replaced); (3) MLR (mean number of
syllables between all pauses); (4) MPT (mean number of seconds for all filled and
unfilled pauses); and (5) TPT (total pause time [filled and unfilled]): expressed
as a percentage of the total time on task). In addition, the Appendix shows that the
post-hoc results revealed significant differences between the V/NV and NB
conditions in relation to the Rate A ($p<.008$), Rate B ($p<.013$) and MPT ($p<.010$)
measures.

In essence, the results obtained by Wolf’s (forthcoming) study indicate that the
fluency of non-advanced Japanese learners of English appears likely to fluctuate in
response to various kinds of listener backchannel responses during L2 English oral
tasks. It is tentatively suggested that non-advanced Japanese learners of English are
particularly sensitive to the presence or absence of listener backchannels during their
talk. Thus, when few backchannels are given, this may serve to reduce their fluency
during oral tasks that require extended periods of production such as narratives.
This study provides support for the backchannel output hypothesis. With a larger
N-size, it seems likely that statistical significance would have also been found,
especially between the V/NV-NV conditions, but perhaps also between the NV-NB
conditions.

6. Summary

As global tourism and other types of international exchanges continue to grow,
greater understanding of factors that might facilitate or impede intercultural
communication is needed. One such factor was suggested to be how individuals
from different sociocultural and linguistic backgrounds use and react to
backchannels which is also of interest to language teachers and researchers. To
demonstrate how sociocultural and linguistic groups can differ in relation to this
aspect of communication, the backchannel behavior of Japanese and American
English speakers was contrasted in terms of the frequency, placement, functions and
types of cues used. Turning to Appraisal Psychology, it was suggested that L2
learners of English from various L1 and sociocultural environments can be expected to have different preferences for receiving backchannels during their talk which operate along the dimension of pleasantness at the schematic level.

Backchannels were described as emotional stimuli that L2 speakers evaluate via their somatic value systems and that may, depending on their preferences, facilitate their fluency. In line with this, I proposed the backchannel output hypothesis which suggests that non-advanced learners of English from various sociocultural and L1 contexts will tend to evaluate verbal and nonverbal English backchannel behavior differently based on their implicit preferences for receiving backchannels, and these evaluations will determine whether or not English backchannel cues facilitate their fluency during oral tasks. It was noted that previous investigations had not considered the influence of listener backchannels on the fluency of L2 learners of English during oral tasks. Wolf’s (forthcoming) study, which found backchannels to facilitate the fluency of non-advanced Japanese learners of English during oral tasks, is a step towards filling this gap in the research literature and provides support for the backchannel output hypothesis.

7. Suggestions for Future Research

For non-advanced Japanese EFL learners, Wolf’s (forthcoming) study suggests that verbal and nonverbal English backchannel cues can facilitate their fluency during oral tasks owing to their strong preference for receiving backchannels during their talk. However, non-advanced learners of English from other L1 and sociocultural contexts, who possess different implicit preferences for receiving backchannels, may not show improved fluency in response to English backchannel cues. For example, Tao and Thompson (1991) noted that Mandarin speakers use backchannels much less frequently than American English speakers do, while Clancy et al. (1996) reported that Mandarin speakers avoid sending backchannel cues during their interlocutors’ speaking turn out of respect for their right to produce their talk undisturbed. Therefore, one might theorize that non-advanced Chinese learners of English will not show improved fluency during tasks in response to English backchannel cues.

Thus, to test the backchannel output hypothesis, it is necessary to investigate the fluency of non-advanced learners of English from different L1 and sociocultural groups in response to different types and amounts of English backchannel cues.
during tasks. Additionally, although production in Wolf’s (forthcoming) study was elicited via oral narrative tasks, future studies might examine how backchannels influence fluency during more open-ended tasks or perhaps even interviews. Through such research efforts, the merits of the backchannel output hypothesis can be more fully known.

References
across cultures (pp.101–113). New York: Prentice Hall.


**Appendix 1** Fluency results for the fourteen Japanese participants in the backchannel conditions

<table>
<thead>
<tr>
<th>Conditions</th>
<th>V/NV M</th>
<th>SD</th>
<th>NV M</th>
<th>SD</th>
<th>NB M</th>
<th>SD</th>
<th>F-value</th>
<th>Sig. V/NV-NV</th>
<th>V/NV-NB</th>
<th>NV-NB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate A</td>
<td>76.46</td>
<td>18.74</td>
<td>70.03</td>
<td>18.09</td>
<td>67.39</td>
<td>17.44</td>
<td>4.387</td>
<td>.023*</td>
<td>.021</td>
<td>.008**</td>
</tr>
<tr>
<td>Rate B</td>
<td>65.73</td>
<td>18.33</td>
<td>59.34</td>
<td>15.93</td>
<td>56.76</td>
<td>16.07</td>
<td>3.982</td>
<td>.031*</td>
<td>.026</td>
<td>.013**</td>
</tr>
<tr>
<td>MLR</td>
<td>5.53</td>
<td>1.76</td>
<td>5.12</td>
<td>1.58</td>
<td>5.00</td>
<td>1.02</td>
<td>1.710</td>
<td>.201</td>
<td>.079</td>
<td>.066</td>
</tr>
<tr>
<td>MPT</td>
<td>1.98</td>
<td>0.44</td>
<td>2.19</td>
<td>0.47</td>
<td>2.42</td>
<td>0.63</td>
<td>5.199</td>
<td>.013*</td>
<td>.061</td>
<td>.010**</td>
</tr>
<tr>
<td>TPT</td>
<td>46.96</td>
<td>11.07</td>
<td>50.61</td>
<td>10.21</td>
<td>53.13</td>
<td>10.36</td>
<td>3.023</td>
<td>.066</td>
<td>.061</td>
<td>.030</td>
</tr>
</tbody>
</table>

Note: *p < .05. Comparisons for locations of significance were tested at the .017 level and the p-values derived from a one-tailed test. **p < .017.