Listener Responses in Interaction:

A Case for Abandoning the Term, Backchannel

Donna T. Fujimoto

会話における聞き手の反応について

－用語「バックチャネル」使用廃止に対する支持

フジモト　ドナ　T.

Abstract

In the review of research on listener behavior in conversation (including such responses as, *Uh huh, Yeah, Mm*), the most common term used is *backchannel*. However, this term is highly problematic. This paper focuses on the problem of terminology, the difficulties of delineating the functions of listener responses, the hazards of transcription, and the pitfalls of transferring research directly from one language and culture to another without taking into account qualitative differences. The article also refers to research on listener behavior that has helped raise the awareness of the importance and complexity of listener phenomena. The paper makes a strong case for abandoning the term *backchannel* and using a more general term like *listener response, listener token* or *reactive token*.

Key words : backchannel, listener response, listener token, reactive token

（Received September 28, 2007）

抄　　録

会話における聞き手の行為（“Uh huh, Yeah, Mm”等）の研究で最もよく使用される用語は「バックチャネル」であるが、これは非常に問題の多い用語である。本論文は、(1)専門用語の問題、(2)聞き手の反応の機能を分類することの難しさ、(3)転写における問題、(4)質的な差異を考慮することなく、ある研究を一つの言語・文化から別の言語・文化へ直接移し替えることに伴う問題、に焦点を当てる。また、聞き手の行為の重要性や複雑性への意識を高めるのに貢献してきた研究についても触れる。本論文は、バックチャネルの代替語として、「リスナー・レスポンス」・「リスナー・トークン」・「リアクティブ・トークン」といったより一般的な用語の使用を強く支持する。

キーワード：バックチャネル、リスナー・レスポンス、リスナー・トークン、リアクティブ・トークン

（2007 年 9 月 28 日受理）

1. 0 Introduction

In the study of interaction between two people, it is usual to designate one party as the speaker and the other as the hearer or listener. Often researchers refer to the speaker as the primary speaker and the listener as the non-primary speaker, reflecting the fact that generally listeners do not remain completely silent when in conversation. They may utter words or sounds at the same time as they are listening to the other person. The main focus of attention, however, has traditionally been on the primary speaker, as that was the source of the substance of the talk, and his/her output could be recorded and subjected to closer scrutiny.

Recognizing the fact that what the listener is doing is crucial to the interaction itself, there have been more studies which focus specifically upon the non-primary speaker. This paper is an attempt to point out various problems in the research so far regarding listener behavior. It looks at the problems of terminology, of determining functions, of deciding on transcription and of transferring terminology and approaches directly from one language and culture to another. A review of the literature shows that this is a rich area which still requires much more investigation.

1. 1 *The Listener Role Contrasted with the Speaker Role*

In typical interactions the listener is not only listening. S/He also sends verbal and nonverbal signals to the speaker. Thus, the listener's role is not an entirely passive one. At any moment during the course of the interaction, the listener can become the speaker. Researchers who have investigated everyday conversations sometimes marvel at how participants in conversations seem to know the intricate rules of turn taking and talk. Ward and Tsukahara (2000) comment, "...there is the mystery of how 'coordination' is achieved—when two people are talking together, their utterances seldom interfere with each other, despite the lack of any fixed protocol for who may speak when." (p. 1178)

In the past research in linguistics has focused almost exclusively on the speaker or on text. Goodwin (1986), Gardner (2001) and others point to the fact that linguistics has traditionally relied on the spoken and written word as the source for data. As Gardner concedes, "This is understandable to the extent that what language users say or write is available and 'out there' for study, unlike listening (or reading), the processes of which are internal, invisible, and not directly accessible to an observer." (p. 1) Modern linguistics has been greatly influenced by Austin (1962) and Searle's (1969) speech act theory where the interpretation of intent and attitude is important. Grice (1968) influenced the field by proposing a view of communication which focuses on intentions and speaker meaning.

Gumperz (1982) focuses on both speaker and hearer. He contends that Gricean pragmatics is based on analysis which is sentence-based and is "concerned with (shared)

presuppositions in the interpretations of intent." (p. 17) For Gumperz, it is not necessary to try to probe the psychological intent of the speaker. Instead he feels it is important to look at how *intent* is understood or interpreted by the *listener*. He states:

> "We assume such interpretation is a function of
> (a) listeners' linguistic knowledge
> (b) contextual presuppositions informed by certain cues, and
> (c) background information brought to bear on the interpretation. (p. 17)

He suggests that conversational cooperation is negotiated with the help of *contextualization conventions* which are signaling cues that help participants interpret what is going on in the interaction.

> One way in which contextualization conventions function is to serve as guide posts for monitoring the progress of conversational interaction. We use our knowledge of grammar and lexicon, along with contextualization conventions and whatever background information we have about settings and participants, to decide what discourse task is being performed and what activity is being signaled, and this provides information about likely communicative goals and outcome. We then build on these predictions to identify the communicative intent that underlies particular utterances (p. 18).

These contextualization conventions are acquired through experience and for the most part are unconscious to the participant.

Like Gumperz the researchers in conversation analysis (CA) also place importance on the listener (Sacks, 1992; Schegloff, 1982; Jefferson, 1984; Goodwin, 1981; Heritage, 1984). These researchers analyzed natural conversation and rather than searching for linguistic rules they sought to uncover sociological patterns revealed by the interactions. Their work has provided valuable insights into the mechanisms of turn taking and the importance of sequential order. Schegloff (1982) argues that when researchers leave out bits of talk and behavior which are not made by the main speaker, they are losing the very essence of the interactivity between the participants. He drew attention to the importance of minimal vocalizations such as '*uh huh,*' '*yeah,*' '*mm hm,*' etc. and contends that discourse in conversation is an *achievement* of both speaker and hearer, "something 'produced' over time, incrementally accomplished, rather than born naturally whole out of the speaker's forehead" (p. 73). This is a significant departure from the more traditional line of linguistic research.

2. 0 Names for Listener Responses

When reviewing the literature which takes the listener into account, we find there is a very wide range of terminology being used for listener responses. These include:

1) backchannel (Yngve, 1970, many others)
2) backchannel feedback (Ward & Tsukahara, 2000)
3) backchannel message (Tottie, 1991)
4) backchannel item (Tottie, 1991)
5) verbal listener response (Dittman & Llewellyn, 1968)
6) accompaniment signal (Kendon, 1967)
7) minimal vocalization (Schlegoff, 1982)
8) affirmative response (Hirschman, 1994)
9) acknowledge act (Sinclair & Coulthard, 1975)
10) hearer signal (Bublitz, 1988)
11) reactive token (Clancy, Thompson, Suzuki & Tao, 1996; Mori, 1999, 2004)
12) minimal feedback (Holmes, 1997)
13) minimal response (Zimmerman & West, 1975; Maltz & Borker, 1982; Young & Miller, 2004)
14) minimal response token (Young & Miller, 2004)
15) minimal token (Mori, 2004)
16) receipt token (Heritage, 1984)
17) continuer (Schegloff, 1982)
18) continuer token (Young & Miller, 2004)
19) acknowledgement token (Jefferson, 1984)
20) carry-on signal (Stenstrom, 1987)
21) marginal word (Dubois, Schuetze-Coburn, Cumming, & Paolino, 1993)
22) interjectory utterance (Imaishi, 1994)
23) newsmarker (Heritage, 1984; Gardner, 2001)
24) *aizuchi* (for interactions in Japanese) (Maynard, 1986; LoCastro, 1987)

Fries (1952) was the first to write about single utterances in English conversation such as "*Yes,*" "*Uh huh,*" "*Yeah,*" "*I see,*" "*Oh.*" He noted that they are used by the listener to show continued attention. It took two more decades before Yngve (1970) drew attention again to these utterances and coined the term *backchannel*.

> 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 back channel, over which the person who has the turn receives short messages such as *yes*, and *uh-huh* without relinquishing the turn (p. 568).

The implication here is that the speaker is using the primary or *main* channel while the listener uses a lesser channel, the *back* channel. For Yngve, backchannels included all utterances that simply show recipiency or listenership. Oreström (1983) categorized these listener utterances as either *speaking turns* or *back-channel* items. These back-channel items could be both lexical and nonlexical responses. Such responses have "special functions where the listener informs the speaker that his message has been received, understood, agreed to and/or has caused a certain effect" (p. 23).

Duncan (1974) and Duncan and Fiske (1977) broadened the term backchannel to include sentence completions, requests for clarification, brief restatements and non-verbal responses such as head nods and shakes. They propose a signal-based approach to the study of interaction, i.e., that there is a turn-taking mechanism which operates through the use of cues, i.e., turn-yielding cues and turn-continuing cues. They investigated not just vocalized responses but also intonation, para-language, and body motion. In their research they carried out statistical measurements to see what the relationship may be between these nonlexical features, syntax and their overall distribution in relation to turn changes.

In contrast, Sacks, Schegloff and Jefferson (1974) suggest that turn-taking in conversation is governed by a set of rules. Their approach to understanding interaction, known as conversation analysis (CA) and referred to previously, looks for recurrent patterns within the talk. One basic structure in conversation is the adjacency pair, which is a set of two utterances made by two different speakers where the first pair part is followed by its second pair part. The basic unit of turns are referred to as *turn construction units* (TCU), which can be sentential, clausal, phrasal and lexical. Change of turns occur at *transition relevance places* (TRP). Turn taking thus involves "a basic set of rules governing turn construction, providing for the allocation of a next turn to one party, and coordinating transfer so as to minimize gap and overlap" (p. 12).

Because of the conversation analysis principle that no detail of naturally occurring conversation should be overlooked, no matter how irrelevant, uninteresting or accidental it may appear, it is not surprising that CA researchers have contributed to the list of listener responses. Sacks (1971) is credited with coining the term, *continuer*, which was further elaborated upon by Schegloff (1982), and later Jefferson (1984) suggested the term *acknowledgement tokens*.

2. 1 *Problems with Terminology*

For the researcher one of the first problems is deciding which term to use of the long list given in the previous section. It seems that the term *backchannel* is the most widely cited, and it also has entered the vocabulary of the layperson as well. Unfortunately, all the behaviors of the listener have been thrown haphazardly into one overgeneralized term, and it thus loses it value as a research tool.

It appears that many researchers have opted to use the term, backchannel, and have then proceeded to modify its definition to suit their particular research interest. For example, some researchers acknowledge that backchannels can include nonverbal responses, but because of the nature of the research project or the lack of video recordings, they restrict their focus to only audible vocalizations. Tottie (1991) uses backchannel more generically, but then adopts a more extended definition, *backchannel message* and goes on to go into further specificity: "When it is necessary to speak of a particular linguistic form functioning either as a backchannel or as part of one, I shall use the term '*backchannel item*" (p. 256). If the term *backchannel* had been an adequate term in the first place, Tottie would have been spared this trouble!

The most serious problem with the terminology is the fact that researchers are mixing together the name of an entity and at the same time naming it for its functions. Perhaps it is due to the fact that in the past so little attention had been paid to the listener, that when attention was drawn to listener behavior, terminology became hopelessly lumped together. This means that it is not always easy to compare research results until one confirms how the researchers have defined their terms. Some take the time and effort to explain, while others simply use the most convenient term.

To help make sense out of all these terms, let us compile a list of all the possible functions of the generic listener response:

1. continuer—keeps the floor open for the current speaker to continue speaking
2. acknowledgment—shows one has heard the current speaker
3. newsmarker—marks what the speaker has said as newsworthy in some way
4. change of activity—marks the transition for a new activity or topic
5. assessment—evaluates the talk of the current speaker
6. claim understanding—shows one has comprehended the speaker
7. agreement—shows support of what the speaker has said
8. disagreement—shows non-support of what the speaker has said
9. clarification—checks to make sure s/he has heard correctly
10. signal of confirmation—shows one has received specific information
11. interest or attentive signal—displays interest and engagement in what the speaker has said

12. collaborative finish–the listener finishes the speaker's utterance
13. emotional response–expresses an emotional reaction to the speaker

This is not meant to be an exhaustive list of possible functions. They are listed here to illustrate the variety of uses used by researchers. With the sheer number of functions, it is quite probable that any term which tries to encompass all of them will only serve to obscure the differences between them––differences which could be crucial to understanding an interaction.

Drummond and Hopper (1993) seem to agree:

> The failure....to distinguish between different classes of back-channels and the consequences they may have for speakership incipiency has made the back-channels category a hodgepodge—though the concept itself captures a basic intuition about brief turns. The concept remains widely cited, but evidence for its usefulness is thin and undifferentiated (p. 162).

The researcher is also faced with another serious problem. Some of the functions given above are very difficult to identify with 100% accuracy. How does one distinguish between, for example, (2) acknowledgement and (11) signaling interest? Or between (6) claim of understanding and (7) showing agreement? The response itself can take any number of possible forms––it could be a lexical item, a nonlexical vocalization or even a nonverbal response such as a head movement or facial expression. There is no easy one-to-one correspondence with the token and the function.

Let us turn again to the conversation analysts. Sacks (1971) was among the first to discuss response tokens in the 1960s and 1970s, and he observed that *Uh huhs* were highly likely to occur at grammatical completion points. Later Sacks (1992) drew attention to the fact that *Uh huh, Mm hm* and *Yes* often occur when a speaker is telling a story, and these utterances were ways for the listener to display recognition that the story is still continuing (p. 766). As mentioned earlier, he was the first to use the term *continuer* for this phenomenon.

It is significant that Sacks and Schegloff and Jefferson (1974) refrained from adopting the generalized term, backchannels, which others have so quickly adopted. While many other researchers used their intuition or made assumptions about listener responses, the conversation analysts held to the principle that only after extensive analysis of naturally occurring conversation is it possible to discover what is really going on. For example, Sacks in his early lectures (1992) stated, "it would be difficult to say that 'uh huh' exhibits understanding" (p. 746). Attributing meaning or intention to a participant is not part of conversation analysis.

Schegloff (1982) later expanded on the argument saying that *Uh huh* does NOT display

'attention,' 'interest' or 'understanding' because it does not have a semantic component indicating such is the case. He proceeded to explain the function of *Uh huh* through CA's mechanism of *repair*. When there is an absence of a move to repair (note that a repair can be made at any point in an interaction), this is an indication that there is no problem of hearing or understanding on the part of the hearer. Thus, when the hearer says *Uh huh*, this is indicating that the interaction is going smoothly, and there is no claim of a lack of understanding. This is qualitatively different than stating that *Uh huh* shows understanding. Schegloff cautions that when listener responses are taken in aggregate, they completely lose their ability to uncover what is actually happening in a conversation. Thus, from the CA perspective, the list of functions given above would include only 1) continuer and 2) acknowledgement.

2. 2 *Other Possible Classifications*

In attempting to get a better grasp on listener responses, let us try to make some possible differentiations. One clear division is between vocalized and non-vocalized responses, i.e., nonverbal behavior. While vocalized data can be transcribed, nonverbal behavior is much harder to capture and describe. Typical nonverbal behaviors which are considered to be listener responses include head nods and head shakes from side to side. Brunner (1979) analyzed the incipience of smiles and claims that smiles can be considered a kind of backchannel because they often occur in similar locations as other verbal backchannels. Although largely ignored in the past, with advances in technology, more researchers are investigating nonverbal behavior as part of normal interactions.

Then there is the distinction between lexical items and non-lexical items. For example, "yeah" would fall into the first category, but "yaa" would be in the latter group. Ward and Tsukahara (2000) make an interesting comment about non-words. They suggest that noises such as "grunts" could be viewed as vocalizations "custom-made for the occasion, with each of the various elements of the pronunciation of vocalization contributing some element of meaning...that is, sound symbolism may be present" (p. 1200). Perhaps in the past, researchers might have simply discarded such grunts; now they are prepared to find a place for them even if it is a one-time occurrence! Several researchers have pointed out that laughs can function as a type of backchannel (Maynard, 1986; Tottie, 1991; Gardner, 2001). Different kinds of laughter also convey information in an ongoing conversation. Sighs and audible intakes of breath can also be meaningful. Even some coughs and sniffs act like backchannel feedback (Ward & Tsukahara, 2000).

Another distinction often made is that between a listener response which constitutes a turn and one that does not. For many researchers if a listener response functions as a legitimate turn, this will disqualify it as a backchannel. For example, Ward and Tsukuhara (2000) state

that backchannel feedback is sometimes defined as those utterances which "do not take the floor", and/or "are not full turns." Many refer to backchannels as "short utterances," even though as Maynard (1986) points out, Yngve (1970) himself, the originator of the term, backchannel, included longer questions and statements as backchannels, such as, "Oh I can't believe it" and "You've started writing it then—your question?" (p. 574). Many researchers today would not agree with Yngve on these examples. Tottie (1991) and White (1989) note problems of judgment of whether an utterance would be considered a backchannel or a turn especially when the utterance is longer or complex. To decide if an utterance is a backchannel will depend, Tottie contends, on what follows it. If the utterance invokes a response from the speaker, then it should be termed a turn and not a backchannel. Sometimes what began as a backchannel may end up becoming a turn, when the speaker falls silent and does not continue. It seems the line between a backchannel and a turn is very faint and ephemeral.

Tao and Thompson (1991) state that in order to understand a backchannel it is necessary to utilize the concept of "primary speakership." "That is, backchannels will normally not disrupt the primary speaker's speakership" (p. 210). Maynard (1986) also sees the conversational turn as essential to the definition of the backchannel. "...we define the listener's back channel in the context of 'turn'--whether the listener's utterance is given during the other participant's speaking turn or not" (p. 1084). Schegloff (1982) (who as we noted earlier, deliberately does not use the term backchannel) examines '*uh huh*' and other minimal vocalizations in terms of rule-governed turn-taking sequential behavior. He notes that sometimes the listener has the chance to begin to speak, but does not and instead uses vocalizations such as '*uh huh*' and '*yeah.*' This allows the speaker to continue the turn, and the vocalization is thus referred to as a 'continuer.' Schegloff shows how it is structurally relevant for a listener to display understanding through the '*uh huh*' and '*yeah*' at the time that "an extended turn is underway, and to show their intention to pass the opportunity to take a turn at talk that they might otherwise initiate at that point" (p. 81).

Tottie (1991) suggests that backchannels can serve a supportive function or a regulative function. Supportive functions could be those showing agreement, or understanding, while regulative functions would be continuers, or change of activity tokens. This may not be a very useful classification since Tottie notes that there are other researchers like Nordenstam (1987) who feel that backchannels have both functions at the same time.

To summarize, we have looked at the range of listener response in terms of:

1) form vs function
2) verbal vs nonverbal
3) lexical vs nonlexical
4) turn vs. backchannel

5) supportive function vs regulative function

3. 0 Other Perspectives on Listener Responses

There have been other researchers who have investigated listener responses in greater detail. Goodwin and Goodwin (1987) have focused on assessments. They make a distinction between response tokens, such as continuers, and assessments. Assessments take a stance on what has been said, however, continuers do not take a stance, instead responding to what has been said as a preliminary to another unit of talk. Continuers are only recipient actions, while assessments can be accomplished by either a recipient or a speaker. An assessment may occur at the end of an extended turn, while a continuer occurs during the extended turn. As Gardner (2001) notes,

> We are then, according to Goodwin, dealing with two structurally different ways of dealing with another's talk. Assessments comment on what another has said, without treating it as a preliminary to something else, and continuers...treat the talk to which they are responding as an emerging element in a larger, as yet incomplete structure" (p. 6).

Stubbe (1998) claims that listener responses help to set up a general "sharing of a frame of reference" between the speaker and the listener. These brief responses can show many different kinds of meanings regarding what has been said and even the listener's attitude towards the speaker. Stubbe suggests a continuum for listener responses with one end indicating low involvement and neutral affect (e.g. relative indifference or simple affirmation) and the other end being high involvement and positive affect (e.g. enthusiastic interest and agreement).

Gardner (2001) points out, however, that in delineating what makes an utterance neutral, positive or negative, it is necessary to also take into account important characteristics such as intonation contour and prosodic shape. "...a supposedly neutral, minimal token such as *Mm* that is...lengthened, uttered loudly, and rises to high pitch can, given an appropriate placement in a sequence, convey a high involvement, positive affect message" (p. 19).

Müller (1996) also agrees that great attention should be paid to prosody when investigating brief responses. He looks closely at continuers and acknowledgement tokens, at their prosodic characteristics and their exact sequential locations. He points out that in contrast to disaffiliative tokens, affiliative tokens have more variation in intonation, lexical possibilities and range of length.

In his very thorough investigation, Gardner (2001) uses only the first five entries on the list of functions given above. For him response tokens are: 1) continuers, 2) acknowledgement

tokens, 3) newsmarkers and 4) change-of-activity tokens, while 5) assessments form a different category. All give information about the talk that has already unfolded and give indications of talk projected in the forthcoming talk. Gardner's work illustrates how these brief utterances that the listener makes have the power to influence the ongoing talk. As he points out, the primary function of response tokens is "not to make reference to the world but to provide some information on the course the talk is taking" (p. 14).

3. 1 *Phonetic Renderings of Listener Talk*

There is an extremely wide range of phonetic renderings of listener utterances within the research literature. The following is just a sampling:

Tottie's (1991) list include: YEAH, MHM, HM, RIGHT, UNHHUNH, UHUH, SURE, YES, HUHN, YES, NO, REALLY, and OH. Then there is a more detailed list of the variants of MHM which is divided into three groups: simple, double and complex.

Simple: *mhm*, *m=hm*, *m=hm=*, *mhm=*

Double: *m=hm mhm*, *mhm mhm*, *mhm=mhm*, *mhm mhm=*, *mhm mhm mhm*

Complex: *mhm hm=*, *m=hm . u=nhhunh*, *m=hm mhm yea=h*, *m=hm .. mhm*, *gee*

(p. 267)

In the same study there are long lists of variants of both "yeah" and "yes" which will not be reproduced here. (see pp. 267-268)

Clancy, Thompson, Suzuki and Tao (1996) show the following:

English: hm, huh, oh, mhm, uh huh

Japanese: un (u=n, unun, etc.), a=, ee, ha=, ho=, hu=n, he=

Mandarin: uhm, a, ao, ai=, en=, eh, hum, mhm=/mh

(p. 359)

4. 0 A Closer Look at Eight Response Tokens

Just this brief look points to the sheer number of possible renderings of listener utterances. They are multi-functional and can be very complex. However, as Gardner (2001) points out "Whilst response tokens are highly variable, even unstable, they do have identifiable core forms and core uses" (p. 8). On this premise he proceeds to examine eight response tokens.

(1) *Mm hm*

(2) *Uh huh*

Continuers: The protypical continuers are *Mmhm* and *Uh huh*. Research supports the claim that these two utterances are used in the same way. Investigations seem to indicate that these are more common in American English than British or Australian English (p. 27). These two utterances generally occur alone without additional brief utterances or more substantial

topical utterances. Thus, they are different from *Yeah* and *Mm*, which can be accompanied by other speech. *Yeah* and *Mm* can also function as continuers, but are more commonly used as acknowledgement tokens. Compared to *Mm hm* and *Uh huh*, *Yeah* shows a greater degree of speaker incipiency (Drummond & Hopper, 1993; Jefferson, 1984), meaning that there is a greater chance for the listener to follow *Yeah* with further substantive talk. According to their research, *Yeah* signals bids for a turn 50% of the time, while *Mm hm* does so 5% and *Uh huh*, 4% of the time.

(3) *Yeah*

(4) *Mm*

Acknowledgement tokens: The most common are *Yeah* and *Mm* which have a falling intonation. When they end with a rising intonation pattern, they tend to act as continuers. *Yeah* is the most frequent acknowledgement used in English and is, in fact, the most common of all the listener response tokens in English conversation. They differ from continuers because they do not give the floor to the speaker but claim simply to have adequately received the prior turn. Thus, unlike the continuer, which refers to the present talk, the acknowlegement token is retrospective. Drummond and Hopper (1993) suggest that *Yeah*, more so than *Uh huh* and *Mm hm*, may signal that the listener is about to shift out of the recipient role and may in fact be projecting forward to a speaking turn. This seems to support Jefferson (1993) who found that *Yeah* was frequently followed by a shift in topic, what she referred to as a *preshift token*.

Compared to *Yeah*, *Mm* is generally seen as weaker and shows less involvement. According to Gardner (2001) "*Mm*...is weak and minimal, arguably the most minimal of all vocalisations in conversation. *Mm* can be seen as a non-intrusive, reserved response to a delicate topic" (p. 32). Gardner shows many instances where uses of particular utterances showed skill in the management of interpersonal interaction.

(5) *Oh*

(6) *Right*

Newsmarker: These responses mark the prior turn as newsworthy to the listener. The core tokens are: *Oh*, *Right*, *Really* (and sometimes short questions like, *Did they?*). Jefferson (1978) notes that *Oh* sometimes acts as a *disjunct marker*, when the listener suddenly remembers something which may not be directly related to the prior talk. She also claims that *Oh* may mark a *noticing*. Heritage (1998) found that *Oh* is usually followed by further talk although it can also occur alone. *Ohs* are "momentary indicators of a speaker's state of mind" (p. 43). Heritage (1998) also found that *Oh* can often occur just prior to a question.

Gardner (2001) notes that there are differences in the usage of *Right* between Americans, on the one hand, and the British and Australians on the other. The most frequent use of *Right* in American English is as an agreement marker, not as a newsmarker. The newsmarker Right

is "used to accomplish a recognition on the part of the utterer that the unit of talk to which it is responding, or an idea from that unit of talk, has been understood to be connected to another unit from earlier in the talk" (p. 47).

(7) *OK*

(8) *Alright*

Change of activity: The most common tokens are *OK* and *Alright*. These signal that there will be a change of topic or a change of activity. Schegloff and Sacks (1973) refer to the place of occurrence of *Okay* and *Alright* as a *pre-closing environment*. *Okay* marks junctures in the talk where a shift is proposed and subsequently negotiated. Thus, it is not uncommon for *Okay* to be followed by further talk on the part of its speaker. While *Alright* has the same function as *Okay*, it is possible that *Alright* may be a stronger signal for a shift and may be given previous to a more major transition than *Okay*.

4. 1 *Five Types of Mms*

Gardner (2001) has made a very positive contribution towards raising awareness of the work that listeners do. Although the utterances listeners make are semantically weak and sometimes empty, they play a crucial role in the co-construction of the ongoing talk. When isolated, it is often difficult to determine what these utterances mean because they are so flexible, multi-functional, and hard to pin down because in one context it may take on a meaning which is quite different from that in another context. Gardner delineates five different types of non-response *Mms*. 1) the lapse terminator; 2) the degustatory token; 3) the hesitation marker; 4) the repair initiator; and 5) the answering token. He also gives a detailed analysis of three response tokens: *Mm* as 1) the weak acknowledgement; 2) the continuer and 3) the assessment.

5. 0 The Problems with Transcription

It is never easy to accurately capture naturalistic interaction. If visual and auditory recordings are made, a major hurdle is deciding what type of transcription coding to use in order to render the recordings in a written form which others can make sense of. The researcher has to select an approach to transcription which will conform to the research stance of the study, yet at the same time be aware that the transcription itself can mask or illuminate important elements.

The transcription of nonlexical vocalizations or non-words is problematic. How can a researcher accurately represent such sounds as: *Mm*, *umm*, *mn-hm*, *uh-huh*, *uhh-hunh*, *yeh*, *yeah*, etc? Researchers have arbitrarily come up with their own set of orthographic representations, but they will naturally differ from person to person. (See the previous section.) Dictionaries are of no help as only a few minimal utterances have ever been

candidates for dictionary entries. Transcription methods books (eg. Edwards & Lampert, 1993) encourage researchers to make their transcripts readable by keeping in mind the expectations of readers (p. 6). This is why transcriptions have often used similar representations to those of novelists who have managed to capture the flavor of talk in the real world.

Intonation can convey significant meaning, and this must also be faithfully rendered on the page. Many factors such as, pitch, volume, stress, duration and prosodic contours hold meaningful information. Koiso, Horiuchi, Tutiya, Ichikawa and Den (1998) and Ward and Tsukahara (2000) investigated backchannels and prosodic features. The use of the computer has greatly aided their ability to share their research results. Along with the pitch contours of both participants, researchers can also show the transcription and other comments on the context on the same page.

According to Zimmerman (1993) coding can obscure differences in prosodic shape and intonation contours and hide the importance of the sequential environment. Coding should be done only after close observation of the details of talk and of the environment and should NOT be based on the analyst's intuitions. The data must speak for itself! Coding holds the danger of obscuring complexity, and sometimes it encourages generalizations that are blind to subtleties.

While many researchers have been conscientious in carefully describing phonetic and intonation contours, no two people will agree 100% of the time. Using only words to describe pitch, volume, intonation and prosodic changes in an utterance is cumbersome and problematic. It can make it difficult or impossible to compare research results simply because of the coding conventions used. Ask even well-trained researchers what the difference between the following are,

Mm and *M*

Hmm and *Umm*

Mm hm and *Mhm*

and the chances of getting widely variant answers is highly likely. Matters are complicated further by the fact that cultural, regional, gender and individual differences are also unavoidably at play.

6. 0 Listener Behavior Across Language and Culture

There has been an increasing interest in research of listener talk in other languages and cultures. Most notably there has been a significant response by Japanese researchers, perhaps because even the Japanese layperson knows a rough Japanese equivalent for the term, backchannel, i.e., *aizuchi*. (See Number 24 in the list of names for listener talk in section 2.0.) Maynard (1986) and Yamada (1992) found that Japanese tend to make listener-type responses, both verbal and nonverbal, roughly twice as often as Americans. White (1989)

found that "Japanese appear to have a higher overall baseline for backchanneling in daily conversation" (p. 67). She suggests that it is the speakers who provide more opportunities for backchannels, and at the same time the listeners use more of those openings.

Not only that but the Japanese language itself may be a factor in eliciting more backchannels. Ward and Tsukahara (2000) suggest that a low pitch may signal backchannel responses, and if that is the case, then the occurrence of clause connectives such as *kara* (because), *te* (and) and *kedo* (but) at these low pitch areas are also opportunity spaces for backchannels. The same researchers also found that sentence-final particles, such as *ne,* (y' know) were frequently present at low pitch areas as well. This seems reasonable since *ne* often functions as a way to 'seek agreement' or 'invite collaboration' (Cook, 1992). Mizutani (1988) supports the claim that backchannels occur near conjunctive particles.

However, there is a serious problem here. Many of the researchers have not been very exact with their definitions, and they have lumped together many of the items found in the list in section 2.0. This is very unfortunate. It is even more disappointing that many have quickly taken up the Japanese word, *aizuchi*, and then used it interchangeably with *backchannel*. This act has effectively masked some very interesting qualitative differences between the listener behavior in English and that in Japanese.

The word, *aizuchi*, is a combination of the verb, *au*, meaning doing something together, and the word, *tsuchi*, meaning hammer. The image conjured up, explains Yamada (1997), is that of "two blacksmiths hammering away in rhythmic ensemble. A picture of Japanese rhythm in action, listeners chime in, '*n*, *n*, *n*,' (*uh huh*, *uh huh*, *uh huh*) to provide constant accompaniment for a speaker's talk" (p. 96). Maynard (1997) points out that there is a rhythmic synchrony of two participants in a conversation in Japanese. She says,

> Repetitious head movement contributes to the rhythm by beating the tempo of the conversation. In the data examined, the speed of each head movement seems to match the overall speed of conversation: fast-paced conversation is accompanied by fast head movement, slower conversation by slower head movement. When head movements appear in groups, they do not occur randomly, but are distributed in such a way as to be synchronized with the tempo of the talk" (p. 151).

She likens the conversation then as a synchronized dance which is a demonstration of "constant and consistent empathy-building on the part of both participants" (p. 151). Maynard found instances of mutual headbobbing in her data of Japanese participants, but none for those for the American participants, which she says is the most striking difference between the English and Japanese data. Despite the fact that Maynard is very aware of these qualitative differences in Japanese conversation, she persists in using the word *backchannel* even when

referring to Japanese interaction.

There are other ways that Japanese interaction may be qualitatively different from interaction in English. For example, Mizutani and Mizutani (1987) claim that "A lack of *aizuchi* sends a message to the speaker that the listener has not understood or does not want to continue the conversation" (p. 21). In a similar interaction in English, there is a greater likelihood that there would be no such expectation, and the speaker may suspect a problem if there is a change in eye gaze perhaps or a verbal response bringing up that problem.

Unlike the backchannel, *aizuchi* has another unique function. Matsuda (1988) suggests the concept of *ma wo motaseru*. *Ma* is an important Japanese space-time concept which indicates a "meaningful pause, interval, or space...In speech this means that it is the silences between words that also carry meaning and are significant." (Hall, 1983). This suggests that both participants may be using *aizuchi* in such a way that the line between the speaker and the hearer (or the primary and non-primary speaker) becomes blurred or nonexistent. This becomes problematic if one insists on using backchannel as one's terminology. The unfortunate conclusion would likely be totally overlooking this type of interaction.

Mizutani (1988) has coined the term, *kyowa* to capture characteristics of Japanese conversational style, where the sentences are constructed cooperatively by the interactants. This contrasts with the English conversation style, *taiwa* or dialog. It is also interesting to note that when White (1989) failed to find purely linguistic reasons for the higher incidence of backchannels among Japanese than for Americans, she turned to the examination of cultural values. She uses the concept of *omoiyari* which has no equivalent term in English. White writes, "the concept generally refers to the creation and maintenance of smooth and pleasant human interactions....To maintain harmony, unanimity, or mutual understanding, people must be most sensitive to the recipient's point of view and feelings" (p. 67). Thus, she suggests that it is this Japanese ideal which influences hearers to try to "anticipate, understand, and accommodate" their partner.

Taking this kind of approach in investigating phenomena from other cultures and other languages, although difficult, may keep the researcher open to the possibility of discovering heretofore undiscovered elements which play a role in conversation. If White can make use of a Japanese concept with no English equivalent, there is no reason why Japanese researchers cannot retain the Japanese term, *aizuchi*, especially when it has more power to shed light on conversation than the less useful word, backchannel.

7. 0 Conclusion

It is clear then that the term *backchannel* is problematic. Researchers have lumped many different items within this rubric without taking the time to define the specific phenomena that they are focusing upon. In many cases the term *backchannel* has been adopted with a

passing reference to its origins, but then it is modified to fit the needs of the research project. This makes it very difficult to compare studies, and the result is generally that important distinctions and characteristics of conversation are obscured and lost. Thus *backchannel* has outgrown its usefulness for researchers because it encompasses far too many different forms with too many different functions. Rather than clarify, it masks the finer details of what listeners do by overgeneralizing.

Perhaps the biggest problem with the term is that it communicates the idea that what the listener does is of lesser importance than that of the speaker. This fails to take into account the fact that the listener and speaker jointly produce the conversation, and thus one is not lesser than the other. They are simply accomplishing different work at any one given moment. This concept is more aptly conveyed by the Japanese term, *aizuchi*, which stresses mutual participation in talk.

This is not a recommendation, however, to replace *backchannel* with the term *aizuchi*. Far from it because the blanket use of *aizuchi* will lead to the same problems as using *backchannel*, and *aizuchi* has a particular connotation tied to the Japanese style of communication which should not be generalized. This paper has made it quite clear that what listeners do in conversation is complex, and thus a term which tries to encompass all will simply be counterproductive. No single term can possibly cover everything yet have enough explanatory power to be meaningful. It may be best NOT to search for an all-encompassing term but instead to use a more general term. *Minimal vocalizations* has worked adequately enough for Schegloff (1982) and other conversation analysts. Instead of focusing on terminology, these researchers investigate the finer details of what listeners are doing and what these vocalizations mean within the conversation. *Minimal vocalization* does not, however, cover responses which are nonverbal, and, as researchers have shown, these responses are important components of conversation. Thus, it may be best overall to utilize even more general terms, such as *listener response*, *listener token*, or *reactive token*. The researcher is then free to further identify the features and functions of the listener phenomena being studied. By abandoning the term *backchannel*, it is hoped that researchers and practitioners alike can become more aware of the complex work of the listener within conversations.

References

Austin, J. (1962). *How to do things with words*. Oxford: Clarendon.

Brunner, L. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, 37(5), 728-734.

Bublitz, W. (1988). *Supportive fellow-speakers and cooperative conversations*. Amsterdam and Philadelphia: Benjamins.

Clancy, P., Thompson, S., Suzuki, R., & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics*, 26(3), 355-387.

Cook, H. (1992). Meanings of non-referential indexes: A case study of the Japanese sentence-final particle, *ne. Text*, 7, 507-539.

Dittman, A., & Llewellyn, L. (1968). Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology*, 9,(2), 79-84.

Drummond, K., & Hopper, R. (1993). Back channels revisited: Acknowledgement tokens and speakership incipiency. *Research on language and social interaction, 26*(2), vol. 26, 157-177.

Dubois, J., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. Edwards & J. Lampert (Eds.), *Talking data: Transcription and coding in discourse research*(pp. 45-89). Hillsdale, NJ: Lawrence Erlbaum.

Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language and society*, 2, 161-180.

Duncan, S., & Fiske, D. (1977). *Face-to-face interaction: Research, methods, and theory*. Hillsdale, NJ: Lawrence Erlbaum.

Edwards, J., & Lampert, J. (1993). *Talking data: Transcription and coding in discourse research*. Hillsdale, NJ: Lawrence Erlbaum.

Fries, C. (1952). *The structure of English*. New York: Harcourt Brace and Co.

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

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, C. (1986). Between and within: Alternative sequential treatment of continuers and assessments. *Human Studies*, 9, 205-217.

Goodwin, C., & Goodwin, M. (1987). Concurrent operations on talk: Notes on the interactive organization of assessments. *IPRA Papers in Pragmatics*, 1(2),1-54.

Grice, H. P. (1968).Utterer's meaning, sentence meaning and word meaning. *Foundations of Language, 4*, 225-242.

Gumperz, J. (1982). *Language and social identity*. Cambridge: Cambridge University Press.

Hall, E. (1983). *The dance of life*. New York: Random House.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. Atkinson, J. Maxwell & J. Heritage (Eds.), *Structures of social action* (pp. 299-345). Cambridge: Cambridge University Press.

Heritage, J. (1998). Oh-prefaced responses to inquiry. *Language in society, 27*, 291-334).

Hirschman, L. (1994). Female-male differences in conversational interaction. *Language in society, 23*, 427-442.

Holmes, J. (1997). Story-telling in New Zealand women's and men's talk. In R. Wodak (Ed.), *Gender and Discourse* (pp. 263-293). London: Sage.

Imaishi, S. (1994). The use of *aizuchi* in natural Japanese discourse. *Nihongakuho, 13*, Osaka University, Japan, 107-121.

Jefferson, G. (1978). Sequential aspects of storytelling in conversation. In J. Schenkein (Ed.). *Studies in the organization of conversational interaction* (pp. 219-248). New York: Academic Press.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens 'yeah' and 'mm hm.' *Papers in Linguistics, 17*(2), 197-216.

Jefferson, G. (1993). Caveat speaker: Preliminary notes on receipt topic-shift implicature. *Research on Language and Social Interaction, 26*(1), 1-30).

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*, 22-47.

Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A., & Den, Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and Speech 41*, 295-321.

LoCastro, V. (1987). Aizuchi: A Japanese conversational routine. In L. Smith (Ed.), *Discourse Across Cultures: Strategies in World Englishes* (pp. 101-113). New York: Prentice Hall.

Maltz, D., & Borker, R. (1982). A cultural approach to male–female miscommunication. In J. Gumperz (Ed.), *Language and social identity* (pp. 195-216). Cambridge: Cambridge University Press.

Matsuda, Y. (1988). Taiwa nihongo kyouikugaku—aizuchi ni kanren shite, *Nihongogaku, 12*, 59-66.

Maynard, S. (1986). On back-channel behavior in Japanese and English casual conversation. *Linguistics, 24*, 1079-1108.

Maynard, S. (1989). *Japanese conversation*. Norwood, N.J.:Ablex.

Maynard, S. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. *IRAL 35*(1), 37-60.

Mizutani, N. (1988). Discussion on backchannels (in Japanese). *Nihongogaku, 7*, 4-11.

Mizutani, O., & Mizutani, N. (1987). *How to be polite in Japanese*. Tokyo: The Japan Times.

Mori, J. (1999). *Negotiating agreement and disagreement in Japanese*. Amsterdam and Philadelphia, PA: John Benjamins.

Mori, J. (2004). Pursuit of understanding: Rethinking 'negotiation of meaning' in view of projected action. In R. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 157 − 177). London and NY: Continuum.

Müller, F. (1996). Affiliating and disaffiliating with continuers: Prosodic aspects of recipiency. In Couper-Kuhlen and M. Selting (Eds.), *Prosody in conversation* (pp. 131-176). Cambridge: Cambridge University Press.

Nordenstam, K. (1987). Kannlig och manlig samtalsstil, research report, Department of Scandinavian Languages, University of Gothenburg.

Oreström, B. (1983). Turn-taking in English conversation. *Lund studies in English* 66. Lund: Gleerup.

Sacks, H. (1971). Unpublished transcribed lectures, University of California, Irvine. (Transcribed and indexed by G. Jefferson).

Sacks, H. (1992). *Lectures in conversation*. vols. 1 and 2. Oxford: Blackwell.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking

for conversation. *Language, 50*, 696-735.

Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71-93). Washington: Georgetown University Press.

Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*(28), 9-327.

Searle, J. (1969). *Speech acts*. Cambridge: Cambridge University Press.

Sinclair, J., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.

Stenstrom, A. (1987). Carry-on signals in English conversation. In W. Meijs (Ed.), *Corpus linguistics and beyond* (pp. 87-119). Amsterdam: Rodopi.

Stubbe, M. (1998). Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics, 29*, 257-289.

Tao, H. & Thompson, S.A. (1991). English backchannels in Mandarin conversations: A case study of superstratum pragmatic interference. *Journal of Pragmatics, 16*(33), 209-223.

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

Ward, N., & Tsukahara, W. (2000). Prosodic features which cue back-channel response in English and Japanese. *Journal of Pragmatics, 32*(8), 1177-1207.

White, S. (1989). Backchannels across cultures: A study of American and Japanese. *Language in society. 18*(1), 59-76.

Yamada, H. (1992). *American and Japanese business discourse: A comparison of international styles*. Norwood, NJ: Ablex.

Yamada, H. (1997). *Different games, different rules*. New York: Oxford University Press.

Yngve, V. (1970). On getting a word in edgewise. *Papers from the 6th Regional Meeting, Chicago Linguistics Society*. Chicago: Chicago Linguistics Society.

Young, R. F., & Miller, E. R. (2004). Learning as changing participation: Discourse roles in ESL writing conferences. *Modern Language Journal, 88*, 519-535.

Zimmerman, D. (1993). Acknowledgement tokens and speakership incipiency revisited. *Research on Language and Social Interaction, 26*(2), 179-194.

Zimmerman, D., & West, C. (1975). Sex roles, interruptions and silences in conversation. In B. Thorne and N. Henley (Eds.), *Language and sex difference and dominance* (pp. 105-129). Rowley: Newbury House.