

## Addressing the Speech Analyst's Dilemma

- Towards the identification of clausal boundaries in emerging oral exchanges -

Pino CUTRONE

Center for Language Studies, Nagasaki University

カトローニ ピノ  
長崎大学言語教育研究センター

### Abstract

This article addresses a current issue in the field of conversation analysis. Presently, there is little consistency among researchers in the types of speech units used to analyze spoken discourse, which hinders the comparability and replication of findings across studies. This stems from the lack of transparency and authenticity in the literature regarding how speech units are defined and exemplified. Using the specific example of how the primary discourse context attracting backchannels in English, the clausal boundary, is identified in dynamic, emerging and spontaneous oral exchanges, this paper provides a comprehensive and replicable model for future studies to follow. Pedagogical implications are also briefly discussed.

**Key Words:** conversation analysis, emerging speech, speech unit, backchannel, clausal boundary

この論文は、会話分析の分野において現在抱えている問題の一つを取り上げる。現在、研究者の間では、会話を分析するために使用されるスピーチの（構成）単位となるものに一貫性が見られない。そのために取り扱われる研究論文を比較したり利用したりするときの妨げとなっている。この問題は、取り上げられている文献のスピーチの（構成）単位がどのように定義され、またどのように実例として挙げられているのかという観点から見ると、透明性や真正性に欠けていることから生じていると思われる。人と人との会話は、必然的に次々と絶え

間なく変化しながら繰り出される。従って、この論文では、**backchannel** (あいづち) を引き出そうとしている主たる話し手の具体例を挙げながら、その会話のやり取りの中での語と語、あるいは節と節の境界はどのように提示されているのかを明示する。この研究は、広範囲の明確なモデルを提示することを目指す。また、ここでは教授法に関しても論じる。

## 1. Introduction

The analysis of spoken language has long presented challenges to practitioners in various subfields of linguistics (Crookes 1990). There are a number of speech units commonly used in conversation analysis (CA), but the lack of consistency in their use makes it difficult to compare speech analyses across studies. These speech units include Semantic units (such as the Proposition unit, C-unit and Idea unit), Intonational units (such as the Tone unit, Idea unit with a focus on intonation, Utterance unit, and Clause and S-node unit) and Syntactic units (such as the Sentence, Idea unit with a structural focus, T-unit, C-unit and AS-unit). While the writer recognizes that the decision to adopt a particular unit for analysis will, to a large extent, depend on the specific goals of the analyst and the research project, he finds it disconcerting that, too often, these speech units are not comprehensively defined, making it extremely difficult to identify such units in data sets. As Foster, Tonkyn and Wigglesworth (2000: 357) have rightfully pointed out, the following two issues bedevil attempts by speech analysts to make fruitful comparisons and replications:

1. Definitions: ostensibly identical units are either defined in different ways, or not defined at all, or defined in a way that is too simple to deal with real spoken data.
2. Applications: if exemplified at all, definitions are accompanied by one or two citation examples which bear little resemblance to the messy reality of speech transcripts.

Addressing these problems in a specific context, this paper clearly describes, defines and exemplifies the methodology used to identify a particular speech unit in a study involving authentic spoken data. The use of authentic spoken data in this paper (Appendices A and B), containing the problematic and sometimes disorderly discourse features commonly found in natural speech transcripts, is a departure from the tidy,

simplistic and contrived samples of conversation often used in research articles and of limited use to speech analysts. To briefly explain the researcher's focus of study, the next section provides some background information concerning the context of the examples used in this paper. This includes information pertaining to the act of transcribing conversations (2.1), a general introduction to the study of backchannel behavior in CA (2.2), the importance of identifying clausal boundaries in said research field (2.3), and contextual information regarding the conversation that produced the conversational extracts used in this paper (2.4). Subsequently, Section 3 attempts to describe and demystify some of the fundamental challenges analysts face when attempting to identify speech units in emerging, dynamic and spontaneous speech. Section 4, which examines how clausal boundaries are identified in the data, is divided into seven subsections. The first, 4.1, examines the difference between internal and final clause boundaries. The second, 4.2, demonstrates how the researcher deals with ellipted elements and syntactic errors in naturally occurring speech. The third, 4.3, further details the challenges associated with fragmentary discourse commonly found in speech samples. The fourth, 4.4, describes how independent topical noun phrase satellite units can be handled in data sets, while the fifth, 4.5, presents instances of post-clausal text strings functioning as conclusers. The sixth, 4.6, demonstrates how minor utterances and irregular sentences can sometimes function as clause-level concepts, and the seventh, 4.7, explains how interactive aspects of conversation influence the identification of clausal boundaries. Lastly, Section 5 provides some general suggestions for speech analysts to consider moving forward and, subsequently, concludes this paper by offering some pedagogical advice to ELT professionals on how they might incorporate natural spoken data into their classes.

## **2. Context of the examples used in this paper**

### **2.1 Transcriptions**

Most attempts to analyze oral discourse come down to, or at least heavily involve, some type of post hoc analysis of speech transcriptions. As Maynard (1986), Ochs (1979) and Tottie (1991) have long stated, the act of transcribing is somewhat subjective, as the use of any transcription system will foreground various aspects of data and background others. The examples of natural speech used in this paper were taken from a conversational extract produced in the writer's recent study of intercultural communication (IC) in English between Japanese and American participants (Cutrone 2011). Although some contextual information about the study

that produced these transcriptions is provided below to give readers a sense of the dynamics involved in producing the conversation, it is important to note that the objectives, as well as the topic and intended readership, of this current paper are quite different from those of the actual study that produced the transcriptions. That is, while the writer's earlier study rather ambitiously focused on the differing backchannel behavior of cultural groups and how this might inform EFL pedagogy in Japan, this current paper simply attempts to help speech analysts deal with some of the more general problems observed in the field of CA, as stated in Section 1.

## **2.2 The study of backchannel behavior as a field of research**

To provide readers with some general information about the context of the conversations that produced the transcriptions, it is important to note that Cutrone's (2011) multi-faceted research project had numerous objectives. One of them, as mentioned above, was to shed light on the differences in backchannel behavior between Japanese L2 English speakers and American native-English speakers. Backchannels may be defined broadly as the responses and/or reactions that a listener gives to the primary speaker when the primary speaker is speaking (Yngve 1970). In other words, in a dyadic (i.e., paired) conversation, the primary speaker is the one providing the main channel of communication, whilst the primary speaker's interlocutor, the listener, is the one providing backchannel feedback. While the intricacies involved in the various debates within this aspect of pragmatics are clearly beyond the scope of this current paper, curious readers can consult the works of Cutrone (2005, 2010, 2011), Maynard (1997) and White (1989) for a more in-depth account of some of the issues involved in identifying backchannels and how they are distinguished from full speaking turns in conversations. Concerning this field of study in general terms, McCarthy (2002), one of the leading voices in the field of discourse analysis, has recently coined the term "listenership" and has received support from other scholars in identifying this as an important area of study in need of much research attention moving forward (O'Keeffe & Adolphs 2008; O'Keeffe, McCarthy & Carter 2007; Sardegna & Molle 2010; Thonus 2007).

## **2.3 Why are clausal boundaries important to the study of backchannel behavior?**

In studying backchannels, researchers often attempt to decipher whether there are patterns in how and where listeners (also called non-primary speakers) employ these backchannels. Thus, one strategy, which focuses on examining the primary speaker's

speech, is to locate the discourse contexts in the primary speakers' speech which cue backchannels from the listener, i.e., the points in the primary speakers' speech where listeners commonly send backchannels. In several studies involving backchannel behavior, researchers have identified grammatical completion points (in the primary speakers' speech) as places where listener backchannels are commonly found. While it was not entirely clear what Maynard (1987, 1989, 1990, 1997) meant by grammatical completion point, the writer surmises that it was most likely the simultaneous completion of a semantic unit and clausal boundary. From the conversational data their studies produced, Clancy et al. (1996), Cutrone (2005) and White (1989) also observed the clausal boundary to be the single most important discourse context favoring backchannels in English. However, even with their common identification of this discourse context, it is difficult to state with any confidence that their analyses of this discourse context were consistent because their articles did not outline in any great depth the linguistic environs of their transcriptions, nor did they present any examples of how they defined speech units within the problematic areas that are common in dynamic, emerging and spontaneous real-time conversations (Chafe 1988). Therefore, with a focus on the analysis of backchannels in dyadic conversations, the main aim of this paper is to provide a coherent and replicable model for researchers to follow when identifying clausal boundaries. In an attempt to achieve consistency within the field of CA, many of the transcription symbols (see Appendix A) were adopted from the pioneering work of Sacks, Schegloff and Jefferson (1974), and the revised CA conventions of Jefferson (2002).

#### **2.4 Contextual information about the conversation**

The examples of natural spoken discourse used in this paper have been taken from the conversation shown in Appendix B. This intercultural conversation involved two participants who were mildly acquainted: Andrea, a female American exchange student visiting Nagasaki, and Chieko, a female Japanese university student (N.B. Pseudonyms have been used to protect the identity of the participants). Since the conversational participants were both female university students, the researcher was able to control for sociolinguistic variables such as age, gender and social status in the conversation. The conversation, which was video recorded, took place in the researcher's office in Nagasaki Prefecture, Japan, and, once brief instructions were given, only the participants were present in the room. Initial conversational prompts (i.e., involving peer-mentoring) were offered to help stimulate conversation; however,

it was made clear to participants that they were free to talk about anything they liked. The conversation was video recorded for a period of thirty minutes, of which only the middle three minutes of each conversation were included as data to be transcribed (see Appendix B). 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. Lastly, to strengthen the internal reliability of the transcription, the researcher analysed the video recorded conversation with the assistance of a colleague. The colleague, who was trained by the researcher to recognize the transcription conventions used in this study (see Appendix A), assisted by double-checking the researcher's initial transcriptions to ensure that the features of language were accurately presented in the transcriptions.

### **3. Challenges associated with emerging speech**

To understand what is meant by emerging speech, one simply needs to consider the stark contrast between how a person processes spoken as opposed to written language. While the latter, being static and complete, offers the reader a great deal of control and time to reflect, the former does not as listeners are at the mercy of trying to make sense of what the speaker is saying one morphological chunk at a time. Not only is there no definitive way for listeners to predict what comes next, but listeners are faced with the real-time pressure to process and react to what they hear. Unfortunately, too often, conversation analysis is limited to analyzing speech transcripts and the units therein, as static and complete entities. This paper, on the other hand, attempts to better reflect how listener's process language into comprehensible input, and it presents oral exchanges as emerging, dynamic and spontaneous. Therefore, given that location refers to a text string which is developing and open to modification rather than static and complete, categories of structure selected to identify location only reflect the status of the unit at a particular point in the construction of the text (Spelman-Miller 2000). Having identified clausal boundaries as a discourse context of particular interest to the researcher's field of study, the next section attempts to describe the complexities involved in the identification of such units in dynamic, real-time conversations.

## 4. Identifying clausal boundaries

### 4.1 Internal versus final clause boundaries

Considering the challenges presented by emerging speech, Garman's (1990) concept of text unit is particularly useful. This approach begins by "making appeal to verbs as the centers of clausal constructions, treating formal connective devices such as *and*, *if* and relative pronouns as initiating new units, and recognizing unconnected phrases and minor elements as constituting independent units" (Garman 1990: 149). Using text units as a basis, various types of higher order units may be identified as necessary. Within this framework, two types of potential or emergent clause boundaries are identified as discourse contexts favoring backchannels: final and internal clause boundaries (i.e., both with and without a pause). A final clause boundary occurs at points in the primary speakers' speech where the text unit is structurally and semantically acceptable as a clause, but is not bounded by terminal punctuation (Spelman-Miller 2000). These text units are fully meaningful, make complete sense and could end the utterance there (i.e., terminative). In the transcription shown in Appendix B, as well as in the examples below, final clause boundaries are marked by two slashes side by side (/). An internal clause, on the other hand, satisfies the grammatical structure of a clause but not the semantic condition. In other words, internal clauses are marked because the meaning of the clause is not complete, and there is a requirement for the utterance to go on in order for the meaning to be complete (i.e., continuative). A single slash (/) marks the grammatical completion point of an internal clausal boundary. Example 1 below illustrates the difference between the two boundaries.

1. it makes me very nervous/ to to handle that much money//

In this case, an internal clause boundary is located after the main clause *it makes me nervous* because the subordinate clause *to handle that much money* is required to make the utterance fully meaningful. Accordingly, a final clause boundary is marked after the word *money*. Although this analysis was fairly straightforward, there were numerous instances in the data which involved a great deal of discussion between co-transcribers, who were required to reach a consensus on how the transcription conventions would be presented (Appendix A). Clearly, context plays a major role in determining whether or not a potential clause completion point has been reached, and the examples below serve to demonstrate how difficult it is to work with fragmentary

and elliptical data which is common in oral samples, particularly involving the speech of L2 learners. In the original study that produced the transcribed conversation shown in Appendix B, it is important to keep in mind that the identification of boundaries in the Japanese L2 English speakers' speech was relevant to establishing the nature of backchannel placement by their native-English speaking interlocutors in the cross-sectional comparison.

#### 4.2 Dealing with elliptical constructions and syntactic errors

Example 2 below offers a more problematic illustration because it not only contains ellipted elements but also L2 errors in syntax.

2. if you mistake/ very very big problem//

Following Foster, Tonkyn and Wigglesworth (2000), this study recognizes independent sub-clausal units at clause boundaries only when they could be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse. In this way, errors in syntax, which include the erroneous omission of clause elements, function similarly to ellipsis in oral data, and defective clause-functioning strings are, thus, recognized as functioning as clauses in this analysis. As a framework for this analysis, ellipsis was identified according to Quirk et al.'s (1985: 888) description of Situational Ellipsis, which requires the following three criteria: the missing expression is precisely recoverable, the elliptical construction is grammatically defective, and the insertion of the missing words results in a grammatical sentence (with the same meaning as the original sentence). Hence, recovering the ellipted elements from Example 2, this extract is interpreted as a subordinate clause and an independent clause making up the sentence *if you make a mistake, it is a very big problem*. From the context of the situation (see lines 26-32 in Appendix B), as well as a basic understanding of the cause and effect nature of if-clauses, it was determined that the speaker's initial utterance is clearly representative of the final interpretation.

#### 4.3 Working with fragmentary discourse

This section describes instances in the data whereby a string of words was not found to comprise a complete or fully comprehensible clause, as in the following example.



3. so (1.30) so (.65) she said (.90) maybe (.51) \_ (she said) (.) me maybe (1.22)

In Example 3, no clausal boundaries have been marked. While the fragments *she said (.90) maybe* and *\_(she said) (.) me maybe (1.22)* might, at first glance, appear to function as clauses, it is not clear from the context what the speaker was trying to say. The verb *to say* appears to function as a transitive verb, and the word *maybe* appears to function as an adverb and not as a direct object. Ergo, the co-transcribers interpreted *she said* as requiring an obligatory direct object and, thus, as being incomplete.

Directly after the extract illustrated in Example 3, the same speaker uttered the following:

4. i saw<sub>6</sub> (.) i saw every day that brown paper money//

From the speaker's subsequent utterance, the fragments discussed in Example 3 would seem to fit the criteria required of a false start. This is characterized "as an utterance which is begun and then abandoned altogether or reformulated in some way" (Foster, Tonkyn & Wigglesworth 2000: 368). Other dysfluency features found in the data include repetitions and self-corrections. Examples 3 and 4 show instances where repetition is used as *she said* in Example 3 and *i saw* in Example 4 appear to be repeated either as strategies to hold the floor, or to allow time for planning out what to say next.

Example 5 below illustrates an instance where self-correction was used.

5. but you get (.) you make lots of money//

In this extract, the speaker begins with *but you get* and then chooses to reformulate the utterance to *you make*, seemingly to clarify that the person receiving the money had, in fact, earned it. While these types of dysfluency features are prevalent in the data, they have not been shown to be discourse contexts favoring backchannels in previous research and are, thus, not priorities in this analysis. The main reason for distinguishing them here is to demonstrate what does and what does not constitute a clausal boundary, and it is clear that these types of fragments do not constitute a clause. As such, when false starts, repetitions and self-corrections occur in the data, there is nothing demarcating them.

#### 4.4 Independent topical noun phrase satellite units

Another issue concerning the fragmentary nature of oral speech is the problem of identifying independent topical noun phrase satellite units, which are commonly found in speech (Bygate 1988). Such units refer to noun phrases that have a topical association to surrounding speech, as shown in Example 6.

6. you can only use cash// (.) so like real money// paper money//

This extract contains three potential final clausal boundaries. While the first one after the word *cash* is fairly straightforward and its type has been described above, the second two after the word *money* in both cases require some explanation. The speaker's utterances of *real money* and subsequently *paper money* are examples of synonymy and anaphoric references to the word *cash* in the first clause. It seems that the speaker is using these references to clarify precisely what is meant by the term *cash*. In this way, these topical noun phrases or "tails" as Carter and McCarthy (1997: 18) call them, function as clauses in that they take on the constituents of the first clause. These tails can be reconstructed as *you can only use real money//* and *you can only use paper money//* respectively. In some cases, however, a direct relationship from the tail to the preceding clause is not easy to establish. Besides containing a topical association, it is necessary that the noun phrase satellite units occur in close temporal proximity to the preceding clause. Following some of the general principles included in Foster, Tonkyn and Wigglesworth's (2000) analyses, topical noun phrases which are separated from subsequent topical noun phrases by falling intonation (indicated by a  $\downarrow$  sign) and a pause of  $\geq .5$  seconds are not considered to have any topical association. Thus, if Example 6 were altered as shown in Example 7 below, the clausal boundaries would be marked as follows:

7. you can only use cash $\downarrow$ // (.96) so like real money paper money and so they hand it to me//

Since the intonation at the end of the first clause is falling and a pause of  $\geq .5$  seconds occurs, the ensuing noun phrases *so like real money* and *paper money* are not recognized as being connected to the preceding clause and are, hence, not marked as final clause boundaries.

#### 4.5 Post-clausal text strings as concluders

Regarding other post-clausal text strings as text units, this section considers tag questions and other concluders such as *i think, ya know, you see and right* with a rising intonation. When these types of text strings occur after a final clause boundary, they are seen to be connected to the preceding clause and, therefore, the point after which they occur is marked as another final clause boundary, as demonstrated in the following examples:

8. but (.) that was only time// wasn't it?//
9. i'm the only research student// i think//
10. they're boyfriend and girlfriend// ya know//

However, in instances when parenthetical phrases such as *you know* and *i think* are pre-clausal or clause-internal and *kind of* is clause-internal, they are not considered to constitute a clause on their own as shown in the following examples:

11. i think it looks very nice//
12. they're kind of a new group//
13. so she showed (.) ya know the man that guards the gate// (.)

#### 4.6 Minor utterances as clause-level lexical concepts

Another important issue in CA is how minor utterances such as *hello* and *thank you* are interpreted in data sets. While these types of formulae do not conform to the regular patterns of clause structures or to the variations of those structures in the major syntactic classes, they are generally referred to as clause-level lexical concepts or minor utterances/sentences (Evans 2009). Similarly, concerning lengthier utterances, there exist several holophrastic utterances such as *the more, the merrier* and *on with the show*. Concerning the identification of clause boundaries, the class of "Irregular sentences" and "Nonsentences" identified by Quirk et al. (1985: 838-53) are recognized to function as whole sentences and would, thus, satisfy the requirements of a final clause boundary.

#### 4.7 Clausal boundaries in interactive discourse

In the identification of unit boundaries, the connectivity of utterances is not limited to one person's speech. Example 14 below shows how clausal boundaries are interpreted in interactive discourse.

14. Andrea: is it a type of money// (.)  
 Chieko: yeah//

In the dialogue shown in the example above, Chieko's response of *yeah* is interpreted from Andrea's preceding question *is it a type of money*. In this way, based on the preceding sentence, Chieko's utterance of *yeah* is reconstructed to its full form *yeah it is a type of money*, which marks a final clausal boundary.

Another issue related to highly interactive discourse involves how listener interjections are treated. As the researcher's field of study focuses specifically on the listener's perspective, discourse contexts favoring backchannels are identified based solely on the primary speakers' speech and, thus, listener backchannels cannot be considered to complete, or be a part of, the primary speakers' units in any way. Answers to questions (as shown by Chieko's reply in Example 14) are not considered to be backchannels because a question generally imposes a response from the listener, whereas a statement does not; in other words, by definition, backchannels are optional and not required. Examples 15 and 16, which are reconstructions of the more natural and interactive state of how Examples 1 and 2 occurred in the data respectively, illustrate the way backchannels are interspersed in the data. In these examples, the listeners employ the non-word vocalizations *mm* and *A* respectively to signal the primary speaker to continue speaking.

15. Andrea: it makes me very nervous/ (.62) [(to to)] handle that much money// aaa (.53)  
 Chieko: [mm]
16. Chieko: if you (.) mistake/ very = = very big (.) problem//  
 Andrea: = (A) =

Even in cases where backchannels comprise lengthier forms, they continue to not count as separate units nor influence the clausal boundary markings in the primary speakers' speech. Example 17 illustrates how clausal boundaries are not affected by the listener's minor addition backchannel.

17. Andrea: i made a mistake of(1.07) sixty five dollars// (.) (.) which is about nana sen en//  
Chieko: \_ (at work)

However, in studies taking a more macro-level analysis of speech (i.e., not focusing on backchannel behavior), it would make sense to consider listener utterances in terms of how they are connected to what the primary speaker is saying. Thus, as Example 18 demonstrates, the previous utterance could be reinterpreted to include a final clause boundary after the listener's utterance *at work*.

18. Andrea: i made a mistake of(1.07) sixty five dollars// (.) (.) which is about nana sen en//  
Chieko: \_ (at work)//

In this way, the minor addition of *at work* by the listener is seen to complete the primary speaker's utterance to the effect that the listener is conveying *you made a mistake of sixty five dollars at work*.

## 5. Conclusion

As the transcription excerpts above have shown, dealing with spoken data can be a difficult task, fraught with many challenges and choices. This paper has provided a comprehensive account of one researcher's journey in dealing with many of these issues. While some decisions are likely to be based on the specific aims and foci of the research project, others will be left to the discretion of the researcher(s). Regarding the latter scenario, it is imperative that researchers, when writing up and subsequently publishing their work, make concerted efforts to provide comprehensive and transparent definitions and examples of the spoken units used in their research project. Moreover, it is essential that these examples not be of the simple and contrived variety; rather, they should involve the problematic data commonly associated with natural speech samples. To this end, concerning studies involving the analysis of backchannel behavior, or "Listenership" to use McCarthy's (2002) term, this paper has attempted to clearly define and exemplify how the primary discourse context that attracts backchannels, the clausal boundary, can be identified in authentic spoken discourse. It is the writer's modest hope that this increased transparency will lead to more consistency and replication in future studies in this area.

While the contents of this paper are aimed mainly at conversation analysts, some of the general observations and recommendations herein can also be applied to the

domain of language teaching. For instance, somewhat similar to the problems associated with how conversational data is presented in the research literature, ELT materials often include tidy, simplistic and contrived samples of conversation. Consequently, from not having received enough exposure to real-world English, EFL learners are ill-equipped to deal with the subtle, unpredictable and spontaneous nature of English exchanges away from the classroom. To help learners in this regard, the writer encourages ELT professionals to consider innovative methods that will expose their learners to natural spoken discourse, as exemplified in this paper. Recent advances in technology have made it possible for teachers and students to access natural models of spoken English through online corpora, such as the Michigan Corpus of Academic Spoken English (MICASE) and the Vienna-Oxford International Corpus of English (VOICE), to name a few (O’Keeffe, McCarthy & Carter 2007 provide an extensive list of corpora, several of which are freely available online). These corpora provide ELT practitioners with an endless supply of natural spoken texts to use in their EFL classes.

Another means of initiating more authentic models of language into the classroom is for teachers to consider adopting a Task-Based Approach (TBA) to language teaching (Willis 1996). In addition to having a great deal of empirical support in the field of ELT because of its power to motivate learners (Cutrone 2004, Rooney 2000), a TBA stresses the use of natural models of language. That is, the language samples used in a task-based lesson are created naturally by speakers trying to achieve the objectives of a particular task. Conversely, in traditional product-based approaches (as seen in structural and functional/notional syllabi), language sample dialogues are often contrived to model the target structure of the particular unit/lesson, and this, unfortunately, leads to unnatural L2 usage, as learners rely too heavily on specific grammatical structures or set expressions to navigate their way through various situations. Despite its vast potential however, task-based learning has been slow to attract much attention from EFL material writers. Undoubtedly, the traditional product-based approaches are much easier to produce and develop, as EFL professionals may be finding it especially difficult to assess task complexity and sequence/grade tasks in the creation of their task-based syllabi. Ultimately however, since using a TBA means exposing learners to the types of spoken discourse they would experience in the real world, it seems well worth the effort for teachers and researchers alike to iron out the remaining kinks and, thus, make task-based instruction a more viable option in the EFL classroom in the future.

### References

- Chafe, W. (1988). Linking Intonation Units in Spoken English. In J. Haiman and S.A. Thompson, (Eds.). *Clause Combining in Grammar and Discourse* (pp. 1-27). Philadelphia: John Benjamins.
- Clancy, P. M., Thompson, S. A., Suzuki, R. & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics*, 26(1), pp. 355-387.
- Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis. *Applied Linguistics* 11(2): pp. 183-99.
- Cutrone, P. (2004). Designing and implementing a task-based syllabus for a business EFL class. *The Journal of Nagasaki University of Foreign Studies*, 7(1), pp. 57-67.
- Cutrone, P. (2005). A case study examining backchannels in conversations between Japanese-British dyads. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, 24(1), pp. 237-274.
- Cutrone, P. (2010). The backchannel norms of native English speakers: A target for Japanese L2 English learners. *University of Reading Language Studies Working Papers*, 2(1), pp. 28-37.
- Cutrone, P. (2011). Towards the learning of listener responses in the Japanese EFL context. Unpublished PhD Thesis, University of Reading, Reading.
- Foster, P., Tonkyn, A. & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(1), pp. 354-375.
- Garman, M. (1990). *Psycholinguistics*. Cambridge: Cambridge University Press.
- Jefferson, G. (2002). Is “no” an acknowledgment token? Comparing American and British uses of (+)/(-) tokens. *Journal of Pragmatics*, 34(10-11), pp. 1345-1383.
- Maynard, S. K. (1986). On back-channel behavior in Japanese and English casual conversation. *Linguistics*, 24(6), pp. 1079-1108.
- Maynard, S. K. (1987). Interactional functions of a nonverbal sign: Head movement in Japanese dyadic conversations. *Journal of Pragmatics*, 11(1), pp. 589-606.
- Maynard, S. K. (1989). *Japanese Conversation: Self-Contextualization Through Structure and Interactional Management, Volume 35*. Norwood: Ablex.
- Maynard, S. K. (1990). Conversation management in contrast: Listener responses in Japanese and American English. *Journal of Pragmatics*, 14(1), pp. 397-412.
- Maynard, S. K. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. *IRAL*, 35(1), pp. 37-60.

- McCarthy, M. (2002). Good listenership made plain: British and American non-minimal response tokens in everyday conversation. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation* (pp. 49-72). Philadelphia: John Benjamins.
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43-72). New York: Academic Press.
- O’Keeffe, A. & Adolphs, S. (2008). Response tokens in British and Irish discourse: Corpus, context and variational pragmatics. In A. Barron & K. Schneider (Eds.), *Variational Pragmatics* (pp. 69-98). Amsterdam: John Benjamins.
- O’Keeffe, A., McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Rooney, K. (2000). Redesigning non-task-based materials to fit a task-based framework. *The Internet TESL Journal*, 6(12). [online] Available from: <http://www.aitech.ac.jp/~iteslj/Techniques/Rooney-Task-Based.html> (Accessed on March 18, 2012).
- Sacks, H. E., Schegloff, E. & Jefferson, G. (1974). A simple systematics in the organization of turn-taking for conversation. *Language*, 50(1), pp. 696-735.
- Sardegna, V. G. & Molle, D. (2010). Videoconferencing with strangers: Teaching Japanese EFL students verbal backchannel signals and reactive expressions. *Intercultural Pragmatics*, 7(2), pp. 279-310.
- Spelman-Miller, K. (2000). Writing on-line: temporal features of first and second language written text production. Unpublished PhD Thesis, University of Reading, Reading.
- Thonus, T. (2007). Listener responses as a pragmatic source for learners of English. *CATESOL Journal*, 19(1), pp. 132-145.
- Tottie, G. (1991). Conversational style in British and American English: The case of backchannels. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik* (pp. 254-271). London: Longman.
- White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society*, 18(1), pp. 59-76.
- Willis, J. (1985). *A Framework for Task-Based Learning*. Harlow, UK: Longman Addison-Wesley.
- Yngve, V. (1970). On getting a word in edgewise. *Chicago Linguistic Society*, 6(1), pp. 567-578



### **Appendix A: Transcription conventions**

- Listener responses are shown in italics below the primary speaker's talk at the point they occurred in the talk.
- To protect the identity of the participants, pseudonyms are used in the speaker labels on the left side of each transcribed line.
- To not confuse readers with the colons that are used for a different purpose described below, the speaker labels will be followed by a semi colon.
- To further preserve anonymity, pseudographs (i.e., notations in parentheses) will be used in instances where participants' private information such as name, address and/or telephone number has been uttered in the conversation.
- Numbers in parentheses indicate elapsed time in hundredths of seconds of pauses occurring in the conversations. Parentheses with a dot (.) indicate a micropause and/or hesitation under .5 seconds. Pauses are timed using transcription software in this study (Praat Version 5.0.18).
- The equal sign “=” indicates latching - i.e., no interval between the end of a prior piece of talk and the start of a next piece of talk.
- The beginnings of simultaneous speech utterances are marked by placing an opening square bracket at each of the points of overlap, and placing the overlapping talk directly beneath the talk it overlaps.
- Closing square brackets indicate the point at which two simultaneous utterances end.

Metatranscription is shown as follows:

- Empty parentheses ( ) indicate part of the transcription that is unintelligible.
- Words between parentheses indicate the transcribers' conjecture at the words or utterances in the conversation that they are not completely certain of.
- Words between double parentheses may indicate comments and/or features of the audio materials other than actual verbalization.
- L stands for laughter.
- Other than apostrophes, which are used to show contraction between words, punctuation symbols in these transcriptions are not used as regular English punctuation markers indicating grammatical category. While other, non-regular, grammatical functions are shown by symbols such as slashes and double slashes, other punctuation symbols such as question marks and colons are used to indicate prosodic features in these transcriptions.

Nonverbal behavior is shown by the symbols indicated below.

- h stands for audible breathing. ^ stands for vertical head movement (head nod). > stands for horizontal head movement (head shake). S stands for smile. ” indicates that eyebrows are raised. G indicates body or hand gestures.

- In cases where nonverbal behavior occurs concurrently with speech, symbols are placed directly above the speech with which it co-occurs (instances where two types of nonverbal behavior occur simultaneously are shown by underlining them both). Nonverbal behavior that is continuous and occurs for a period longer than 2 seconds will be noted by signaling the beginning and the end of the behavior in parentheses where it occurs in the conversation. N.B. The parentheses containing the symbols below are solely used for separation purposes to make them easily identifiable in the specific examples below. Parentheses are not used in this manner in the transcriptions as they have other specific functions, which have been outlined above.
- A slash ( / ) marks the grammatical completion point of an internal clausal boundary (i.e., a clause which is continuative).
- Two slashes side by side ( // ) mark the grammatical completion point of a final clause boundary (i.e., a clause which is terminative). N.B. A final clause boundary is one that makes complete sense (i.e., fully meaningful) and could end the utterance there. In contrast, an internal clause is one in which the meaning is not complete, and there is a requirement for the utterance to go on in order for the meaning to be complete.
- A question mark ( ? ) at the end of a word and/or utterance indicates a clear rising vocal pitch or intonation (i.e., one that is clearly heard, and is shown to rise by at least 600 Hz using Praat software).
- An inverted question mark ( ¿ ) at the end of a word and/or utterance indicates a clear falling pitch or intonation (i.e., one that is clearly heard and is shown to fall by at least 600 Hz using Praat software).
- A colon ( : ) as in the word “ye:s” indicates the stretching of the sound it follows (i.e., only marked in cases where the stretching was extended greater than .5 seconds).
- A hyphen at the end of an uncompleted word indicates the disfluency of a truncated word. For instance, if the word “word” were truncated, it may be transcribed as “wor-”.
- A part of a word and/or phrase containing CAPITAL letters indicates that it has been said with increased volume and/or more emphatically than the rest of the phrase (i.e., only marked when the highest point of the stressed part of speech is 10 decibels greater than the lowest part of the surrounding parts of speech).
- The underscore sign ( \_ ) indicates that the talk it precedes is low in volume.
- ( ~ ) indicates that the talk which follows is consistent with the person’s regular voice and tone. This symbol is used after low volume talk to indicate the point at which the volume rises back to normal. When a pause occurs after the low volume talk and the talk that follows returns to normal, this symbol is not shown.

**Appendix B: Transcription of conversation**

1. Andrea; is it a is a type of money// (.)

2. Chieko; yeah (.) and ( ) (.) \_ (nantake) (.)

3. Andrea; a (.) rubber bands like a like this// (.) looks like this// = [ok (yeah yeah yeah)]  
 = yeah yeah yeah = [( )]

4. Chieko; so (1.30) so (.65) she said (.90) maybe (.51) \_ (she, said) (.) me maybe (1.22)

5. Chieko; i saw<sub>i</sub> (.) i saw every day/ that [(brown)] money// (2.12) (um) (.)  
 [(yeah)]

6. Chieko; maybe i'm crazy// [ Lh ] Lh =  
 > [(yeah yeah)]

7. Andrea; = that's how i feel// (.51) like people will come up with (.) (uu) (.)

8. Andrea; like (.) ten thousand dollars// (.) (.) ( ) = (.)  
 mmm = ten thousand

9. Andrea; hyaku man en (.88) about (.73) \_ (wait here) (.) (.)  
 ten (.) ( a soka)

10. Chieko; ten [thousand] (.) hyaku man en (.)  
 [( )] mm mhm

11. Andrea; hyaku man en i think (.99) i think right? (.93) no no no (.65)

12. Chieko; no ju man en = (.)  
 = yeah ju man en

13. Andrea; no hyaku no hyaku man = <sup>^</sup> ((G begins)) <sup>^</sup> = hyaku man en. (1.27)  
 = *mm* =

14. Andrea; wait// = [Lh] i'm confused// (.54) \_ ( ) (1.07) (hyaku ) (1.41)  
 = Lh [Lh]

15. Chieko; hey [teachers name] Lh (.)

16. Andrea; [ ah ] hyaku man en = <sup>^</sup> <sup>^</sup> <sup>G</sup> (.) ten thousand dollars about (.)  
 [( )] = *hyaku man men* <sup>^</sup>

17. Andrea; it's about hyaku man en// = <sup>G</sup> (.) and they (.) and it's (.)  
 = <sup>^</sup> \_ (*mm*)

18. Andrea; you cannot use credit cards// (.59) [you] can only use cash// (.)  
<sup>></sup> <sup>></sup> ((G begins))  
 [(*m*)] <sup>^</sup>

19. Andrea; ( ) so like real money paper money. (.) = and so they ha:nd it/ to me// (.)  
 \_ (*mm*) = <sup>^</sup> S

20. Andrea; and i'm like (.) [(ooo oh my god)]// <sup>S</sup> = like h (.) to me i look at that money//  
<sup>G</sup> <sup>^</sup> <sup>^</sup>  
 [(ooo Lh )] Lh =

21. Andrea; and i say [like] (.) kay (.) this money in this hand i could buy a car// (.)  
 [(*L*)] <sup>^</sup> Lh =

22. Andrea; = and i'm like (.50) car (.) [(my)] oh my (.) [(gosh)]// (.) ( ) (.)  
 [(Lh)] [(Lh)]

23. Andrea; cuz i mean i don't have th[at much mon]ey// (.) <sup>></sup> ((G ends)) = like definitely =  
 [ ( \_ Lh )] *yeah* =

24. Chieko; = maybe i think (.) ( <sup>G</sup> ) i want to (.) i want to (1.54)

## Addressing the Speech Analyst's Dilemma

25. Andrea; (you're like [i don't know]// <sup>G</sup> \_ ( ) (.) [( )] (.) [( )] ( ) (.51)  
<sup>G</sup>  
 [ (Lh) ] \_ (Lh) [( )] [( )]

26. Andrea; ((G begins)) aaa (.) but it makes me very nervous/ (.62) [(to to)] handle that much money// aaa (.53) ((G ends))  
 ^  
 [(mm)]

27. Chieko; so (.57) if (.) if you (.) mistake/ (.) very (.) very = = big (.) problem// (.65)  
<sup>G</sup>  
 ^ ^ ^  
 = (A) =

28. Andrea; ((G begins)) [(aa aa a)] big problem// so at the end of the day (.)  
 [( )] \_ (mm) =

29. Andrea; = if if i make a mistake/ i have to count my money?// (.) (.)  
 ^  
 mm

30. Andrea; and (.) say i made a mistake/ of a hundred dollars/ (.65) then (.)  
 ^

31. Andrea; <sup>S</sup> that one hundred dollars/ i have to pay// (1.19) that's why/ it's very like (.57)  
 ((surprised expression))

32. Andrea; ok (.) [you be very] careful// and you count everything//  
 [( )]

33. Andrea; ((G ends)) [like one two three four five] =  
 \_ [( )]

34. Chieko; = that's very nervous job// =

35. Andrea; = <sup>S</sup> yeah it makes me very nervous// = (.) but you get (.)  
 ^ ^ ^  
 = \_ (yeah)

36. Andrea; ((G begins)) you make lots of money// (.65) [ big ] pay check// like your salary is very big// (.77) ((G ends))  
 ^ ^  
 [(mm)]

