The conversational use of reactive tokens in English, Japanese, and Mandarin

Patricia M. Clancy, Sandra A. Thompson, Ryoko Suzuki, Hongyin Tao

Department of Linguistics, University of California at Santa Barbara, Santa Barbara, CA 93106, USA

* Corresponding author. E-mail: sathomps@humanitas.ucsb.edu

National University of Singapore

Received September 1994; revised version April 1995

Abstract

This paper investigates 'Reactive Tokens' in Mandarin Chinese, Japanese, and English. Our definition of 'Reactive Token' (= 'RT') is 'a short utterance produced by an interlocutor who is playing a listener's role during the other interlocutor's speakership'. That is, Reactive Tokens will normally not disrupt the primary speaker's speakership, and do not in themselves claim the floor. Using corpora of conversational interactions from each of the three languages of our study, we distinguish among several types of RTs, and show that the three languages differ in terms of the types of RTs favored, the frequency with which RTs are used in conversation, and the way in which speakers distribute their RTs across conversational units.

1. Introduction

Students of conversational language have noticed for some time that when one speaker projects an extended turn, other speakers may produce small bits of vocal behavior which exhibit an understanding that an extended turn is in progress on the part of the first speaker (Oreström, 1983: 23–25; Sacks et al., 1974; Schegloff, 1982). In this paper we examine these small 'reactive' turns in three languages: Mandarin, Japanese, and American English. Referring to the speaker of the extended turn as the 'primary' speaker, we will suggest that 'reactive' turns play a role in a set of culture-specific communicative strategies that include expectations about how
much input a non-primary speaker will give, and where this input will occur with respect to the primary speaker's turn.

There has developed a profusion of terminology in the last two decades to describe the turns of non-primary speakers. Fries (1952: 49), looking at English conversations, was perhaps the first to group together "those single free utterances ... that have as responses continued attention", including yes, Uh huh, Yeah, I see, Good, Oh, etc. Kendon (1967) called them 'accompaniment signals'. Yngve (1970) proposed the term 'backchannel communication' for all these non-primary turns. Oreström (1983:23), following Yngve, divides "utterances" into "speaking-turns" and "back-channel items", where the latter term includes both lexical and non-lexical "listener responses", representing "rather special functions where the listener informs the speaker that his message has been received, understood, agreed to and/or has caused a certain effect". Duncan (1974) and Duncan and Fiske (1977), however, extended the term 'backchannel' to include sentence completions, requests for clarification, brief statements, and non-verbal responses. Schegloff (1982) marked a critical turning point in the study of 'non-primary' turns, being the first to demonstrate that such turns do not form a single set, but must be analyzed in terms of their interactive functions. What he called 'continuers', in particular, especially uh huh, exhibit the understanding that another turn is still in progress by passing an opportunity to produce a full turn (p. 81). Jefferson (1984) proposed the term 'acknowledgement tokens' for the group of forms in English that includes yeah, mhm, and uh huh, and suggested that functional and sequential distinctions exist among the members of this group. This theme was pursued by Drummond and Hopper (1993a,b), with a response by Zimmerman (1993). Goodwin (1986) made a further contribution to distinguishing among the several types of non-primary vocalizations by proposing an important interactional distinction between continuers and assessments, which evaluate the primary speaker's contribution, such as Wow or Good.

Following several of these scholars, we consider the notion of 'primary speakership' to be critical to the definition of this set of utterances. We also wish to maintain a distinction among several types of utterances produced by non-primary speakers, though for the purposes of this paper, we will propose a different classification based partly on their form and partly on their sequential function, and we will not be comparing the sequential functions of the individual forms within each language, as worthwhile a topic as that would be. We thus term the class of utterances made by a non-primary speaker REACTIVE TOKENS, as indicated in our title. Our definition of 'Reactive Token' (= 'RT') is 'a short utterance produced by an interlocutor who is playing a listener's role during the other interlocutor's speakership'. That is, Reactive Tokens will normally not disrupt the primary speaker's speakership, and do not in themselves claim the floor. If a short utterance served as the second pair part of an adjacency pair (Sacks et al., 1974), for example as an answer to a question or a response to an offer, it was not considered a Reactive Token. In section 4 below we will characterize and exemplify the types of RTs found in our data.

Our study was partially inspired by earlier research showing that 'backchannel' use differs across languages, especially English and Japanese (Iwasaki, 1990; Maynard, 1986, 1987, 1989; White, 1989). These works suggested that Japanese speak-
ers' use of 'backchannel' tokens, or *aizuchi* as they are called in Japanese, is more frequent than that of English speakers. Tao and Thompson (1991), based on the preliminary findings of the research project reported on in this paper, noted that Mandarin speakers use 'backchannels' much less frequently than do English speakers. Mizuno (1988) suggested that Chinese speakers use backchannels less frequently than do Japanese speakers. Liu (1987), in comparing the frequency of backchannels, or *aizuchi*, of Chinese and Japanese speakers reacting to short prepared 'discourses', found that Chinese speakers use *aizuchi* less frequently than do Japanese speakers.

In this study, we do not restrict ourselves to 'backchannels’, since there are other important types of 'non-primary' turn that occur, as we will show, with differing frequencies across languages. Instead, we focus on a more inclusive range of 'non-primary’ turns, or Reactive Tokens, and compare the way speakers from three typologically and genetically diverse languages, American English, Japanese, and Mandarin, use them in everyday interactions. We will provide a cross-linguistic analysis of a variety of Reactive Tokens, and will relate the use of RTs to interactional strategies peculiar to each culture.

We first outline the goals of this study (section 2), and then discuss our data and our methodology (sections 3 and 4). In section 5 we present our findings, and in sections 6 and 7, we discuss the implications of these findings and the conclusions we may draw from them.

2. Goals

The goals of this study are to examine the communicative strategies in each language with respect to culture-specific expectations about the degree of interaction that the non-primary speaker will engage in. There are two related aspects to these expectations, what we might call 'frequency' and 'location'. That is, we will (1) test a prediction arising from the earlier literature regarding frequency of RT use across languages, and (2) demonstrate that the placement of an RT token with respect to the primary speaker's turn is a critical part of the set of expectations. Results of these analyses will provide a basis for proposing a preliminary characterization of the non-primary speaker's role in each language.

3. Data

Our data are all from audiotaped face-to-face ordinary, non-argumentative conversations among friends. By focusing on friends, we hope to eliminate the effects

---

1 From here on, when we say 'English', we are referring to American English.
2 For this project, we have restricted our hypotheses to verbal rather than non-verbal Reactive Tokens. Future research might extend our findings to include an analysis of non-verbal tokens as well (see Maynard, 1989, for findings on Japanese head movements). The Mandarin data are all from native speakers of Mandarin from mainland China.
of differential social status on use of Reactive Tokens; this is, of course, an interesting topic for future study. Tables 1–3 summarize the data from each language (where ‘IU’ refers to ‘intonation units’, to be characterized below).

Table 1
Summary of English data

<table>
<thead>
<tr>
<th>Transcript</th>
<th># of IUs</th>
<th>Speakers</th>
<th>Duration (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farmtalk</td>
<td>204</td>
<td>2 M</td>
<td>6</td>
</tr>
<tr>
<td>Africa</td>
<td>200</td>
<td>2 F, 1 M</td>
<td>3</td>
</tr>
<tr>
<td>Hypochondria</td>
<td>228</td>
<td>2 F, 1 M</td>
<td>4</td>
</tr>
<tr>
<td>Car Sales</td>
<td>204</td>
<td>2 M</td>
<td>3</td>
</tr>
<tr>
<td>Lunch</td>
<td>105</td>
<td>3 F</td>
<td>7</td>
</tr>
<tr>
<td>Dinner</td>
<td>110</td>
<td>2 F, 2 M</td>
<td>7</td>
</tr>
<tr>
<td>Aesthetics</td>
<td>91</td>
<td>1 F, 1 M</td>
<td>7</td>
</tr>
<tr>
<td>Shultz</td>
<td>89</td>
<td>1 F, 4 M</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>1231</td>
<td></td>
<td>44</td>
</tr>
</tbody>
</table>

Table 2
Summary of Mandarin data

<table>
<thead>
<tr>
<th>Transcript</th>
<th># of IUs</th>
<th>Speakers</th>
<th>Duration (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiaoyu</td>
<td>250</td>
<td>2 F, 1 M</td>
<td>6</td>
</tr>
<tr>
<td>Sunday</td>
<td>148</td>
<td>3 F</td>
<td>5</td>
</tr>
<tr>
<td>TKY</td>
<td>137</td>
<td>1 F, 1 M</td>
<td>2</td>
</tr>
<tr>
<td>TKY2</td>
<td>197</td>
<td>1 F, 1 M</td>
<td>4</td>
</tr>
<tr>
<td>HKPR</td>
<td>206</td>
<td>3 M</td>
<td>3</td>
</tr>
<tr>
<td>Shen</td>
<td>126</td>
<td>1 F, 2 M</td>
<td>1</td>
</tr>
<tr>
<td>Thai</td>
<td>112</td>
<td>1 F, 2 M</td>
<td>1</td>
</tr>
<tr>
<td>Tongji</td>
<td>126</td>
<td>2 M</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>1302</td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

Table 3
Summary of Japanese data

<table>
<thead>
<tr>
<th>Transcript</th>
<th># of IUs</th>
<th>Speakers</th>
<th>Duration (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fujikawa</td>
<td>157</td>
<td>1 F, 1 M</td>
<td>3</td>
</tr>
<tr>
<td>Gossip</td>
<td>215</td>
<td>1 F, 1 M</td>
<td>5</td>
</tr>
<tr>
<td>Hamada</td>
<td>237</td>
<td>2 M</td>
<td>5</td>
</tr>
<tr>
<td>Oyama</td>
<td>199</td>
<td>2 M</td>
<td>4</td>
</tr>
<tr>
<td>Girlfriend</td>
<td>121</td>
<td>1 F, 1 M</td>
<td>2</td>
</tr>
<tr>
<td>Party</td>
<td>113</td>
<td>2 F</td>
<td>1</td>
</tr>
<tr>
<td>Surprise</td>
<td>105</td>
<td>2 F</td>
<td>1</td>
</tr>
<tr>
<td>Takando</td>
<td>128</td>
<td>1 M, 1 F</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>1275</td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>
4. Methodology

Our data were transcribed according to the Du Bois et al. (1993) transcription system (see also Du Bois, 1991), which recognizes both turns and intonation units as basic elements of conversational language. The data were then analyzed according to a number of grammatical and interactional coding categories. Here we outline those relevant to the issues we wish to discuss in this paper.

4.1. Speaker change

A speaker change was judged to have occurred at any point at which another speaker took a recognizable turn, whether a full turn or a Reactive Token turn. Laughter turns were counted separately, and will not be included in this discussion.

4.2. Reactive Tokens

We distinguish among several types of Reactive Tokens.

4.2.1. Backchannels

If the Reactive Token is a non-lexical vocalic form, and serves as a ‘continuer’ (Schegloff, 1982), display of interest, or claim of understanding, we consider it a Backchannel. Typical Backchannel forms in each of the three languages found in our data include those found in Table 4.

<table>
<thead>
<tr>
<th>Japanese</th>
<th>English</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>un (u=n, unun, etc.)</td>
<td>hm</td>
<td>uhm</td>
</tr>
<tr>
<td>a=</td>
<td>huh</td>
<td>a</td>
</tr>
<tr>
<td>ee</td>
<td>oh</td>
<td>ao</td>
</tr>
<tr>
<td>ha=</td>
<td>mhm</td>
<td>ai=</td>
</tr>
<tr>
<td>ho=</td>
<td>uh huh</td>
<td>en=</td>
</tr>
<tr>
<td>hu=n</td>
<td></td>
<td>eh</td>
</tr>
<tr>
<td>he=</td>
<td></td>
<td>hum</td>
</tr>
</tbody>
</table>

4.2.2. Reactive Expressions

If the ‘non-primary’ speaker utters a short non-floor-taking lexical phrase or word, we coded that as a Reactive Expression. Typical Reactive Expressions, including
assessments (Goodwin, 1986; Goodwin and Goodwin, 1987, 1992a,b), in the three languages are those found in Table 5.

Table 5
Typical Reactive Expressions in Japanese, English and Mandarin

<table>
<thead>
<tr>
<th>Japanese</th>
<th>English</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>sugoi ‘great’/‘terrible’</td>
<td>oh really/really</td>
<td>zheyang hao ‘such PRT’</td>
</tr>
<tr>
<td>hontoo ‘really’</td>
<td>yeah</td>
<td>jiushi a ‘indeed PRT’</td>
</tr>
<tr>
<td>soo ‘(it is) so’</td>
<td>gee</td>
<td>shi a ‘COP PRT’</td>
</tr>
<tr>
<td>a soo/soo ka ‘(is that) so’</td>
<td>o=kay</td>
<td>dui ‘right’</td>
</tr>
<tr>
<td>ii na= ‘(that’s) nice’</td>
<td>exactly</td>
<td>dui dui/dui dui</td>
</tr>
<tr>
<td>hai ‘yes’</td>
<td>all right</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hell</td>
<td></td>
</tr>
</tbody>
</table>

4.2.3. Collaborative Finishes

When the non-primary speaker finishes a previous speaker’s utterance, we coded it as a Collaborative Finish (Lerner, 1987, 1989, 1991). Of the three languages in the study, we found collaborative finishes only in English and Mandarin. See the Appendix for a list of transcription symbols found in the examples.

(1) English
   A: .. when you say it happens for a reason, 
   .. it’s like, 
   ... it happened to get you off --
   B: .. off my ass.  

(2) Mandarin
   A: ... Yi ge jiao shenme de?
   one CLF call what PRT
   The one, what do you call it?

---

3 We use the following abbreviations for the glosses in our Japanese and Mandarin examples.

2SG = 2nd singular
CLF = classifier
COP = copula
GEN = genitive
NEG = negative
NOM = nominalizer
OBJ = object marker
PRT = final particle
Q = question marker
SUB = subject marker
TOP = topic marker
.. jiao=?
call
it is called
.. haoxiang=,
seem
something like
.. jiao Guoji=... Sh- Shang=ye Xueyuan ba.
call international trade college PRT
International Trade College.
... nei ge xuexiao,
that CLF school
That school
B: ... remer de.
hot PRT
is very popular.
C: .. zui=,
very
very
.. zui remer.
very hot
very popular. (JIAOYU)

4.2.4. Repetitions
If the non-primary speaker reacts by repeating a portion of the speech of the primary speaker, we coded it as a Repetition:

(3) English
A: ... I got everything taken care of.
   I got insurance on it too.
B: ... [how much <X it X>] –
A: ... [under my] name.
   ... eleven hundred a year.
B: ... eleven hundred.
A: ... three hundred [dollars down],
B: ... [that’s cheap] man, (CARSALES)

(4) Japanese
A: .. ototoi–
   the:day:before:yesterday
   the day before yesterday–
   .. kinoo da,
yesterday COP
yesterday
   .. denwa ga atte.
telephone SUB exist
there was a phone call.
B: .. kinoo.
yesterday
Yesterday. (GOSSIP)

(5) Mandarin
B: ... zhongzhuan litou,
intermediate:college inside
Within the vocational schools
hai neng gai bao.
still possible change application
it's O.K. to change your application.
... jieguo,
then,
Then,
haishi mei ren bao.
still NEG person apply
even so nobody is interested in applying (to this type of school).
A: ... ao,
oh,
... mei ren bao gaozhong.
NEG person apply high:school
Even so nobody is interested in applying to such schools.
... jinnian.
this:year
This year.
B: ... mei ren bao gaozhong.
NEG - person apply high:school
Nobody is interested in applying to such schools.
... na=me,
then
So,
... zheyang yi lai,
such once come
in such a situation, (JIAOYU)

4.2.5. Resumptive Openers
'Resumptive Openers' refer to a type of non-lexical element which is used at turn-initial points. These forms would be coded as Backchannels if they weren't followed by full turns. Although in our coding they were treated as a subcategory of Reactive Token, they can be distinguished from both prototypical Reactive Tokens and prototypical contentful turns. No functional claims are intended by the term 'resumptive' in this label.

The characteristics of Resumptive Openers include the following:
1. They are realized in short (typically monosyllabic), non-lexical, vocalic forms.
2. They tend to appear as a separate intonation unit.
3. Normally only short pauses occur after a resumptive opener.
4. They appear at the beginning of a new turn.

(6) English
A: ... How are you doing with the house.
B: ... Oh,
   .. got it all uh... primed,
   .. just about,
   ... except two sides [of it].
A: [Oh you shoot a] primer stuff.       (FARMTALK)

(7) Japanese
A: hyaku nanajuu= ne=,
   hundred seventy PRT
   (she) is a hundred and seventy(-some centimeters) tall
B: ... e=,
   Wow
   ii na=.
   good PRT
   how enviable/(I) envy (her)        (FUJIKAWA)

(8) Mandarin
S: .. Nimen ne ge fangzi,
   2PL that CLF house
   About the housing you have now,
   ... neng bu neng changqi zhu xiaqu ya?
   can NEG can long:term live down PRT
   Can you live there for long?
T: ... (1.0) Na dangran keyi a=.
   that of:course possible PRT
   Sure it's possible.
   ... Ni zhiyao ni yizhi .. zhuce
   2SG just 2SG continuously register
   As long as you register (as a student).
   ... [Na ni jiu keyi].
   then 2SG then possible
   you are fine.
S: [Ai=,
   Oh,
   .. ne ting hao de a=].
   that very good PRT PRT
   that's great.        (TKY)

Resumptive Openers are hybrid in nature: they themselves do not constitute a new
turn, but they are Reactive Tokens that occur at the beginning of a turn. Thus they
differ from the other Reactive Tokens in that their function is to acknowledge the prior turn and commence a new turn, but not to pass a turn-taking opportunity, which is what the other Reactive Tokens may be said to do. Our Resumptive Openers are thus parallel to the ‘acknowledgement tokens’ signalling ‘speakership incipiency’ discussed by Jefferson (1984), Drummond and Hopper (1993a,b), and Zimmerman (1993).

4.2.6. Summary

In this subsection, we have offered a taxonomy of Reactive Tokens which has proved useful in the comparison of the three languages of our study. In the next subsection, as background for our discussion of the location of Reactive Tokens in the three languages, we will discuss our coding of possible conversational completion points.

4.3. Complex Transition Relevance Places

One of our concerns was to determine whether the languages of our study differ in where non-primary speakers tend to place their RTs. In particular, we wanted to know if there are systematic differences as to whether speakers tend to utter RTs at (1) points of possible transition from one speaker to another or (2) during another speaker’s turn. This issue in turn raises the question of how to define points of possible transition from one speaker to another.

Sacks et al.’s (1974) groundbreaking and highly influential paper on the systematic organization of turn-management in ordinary conversation first brought to the attention of students of conversation the issue of conversational units, the units which form the basis for turns. Sacks et al. proposed that turns can be constructed from what they call ‘unit-types’, or ‘turn-constructional units’ (= TCU):

“Unit-types for English include sentential, clausal, phrasal, and lexical constructions. Instances of the unit-types so usable allow a projection of the unit-type under way, and what, roughly, it will take for an instance of that unit-type to be completed. Unit-types lacking the feature of projectability may not be usable in the same way.” (p. 702)

The end of a TCU, then, is what Sacks et al. term a ‘transition-relevance place’, that is, a place where a transition to another speaker might occur. Exactly what defines a TCU is left open in the Sacks et al. paper, but it was assumed that grammatical units play a major role.

Oreström (1983) and Ford and Thompson (to appear) take up Sacks et al.’s challenge to linguists to participate in defining the character of TCUs and the nature of transition-relevance places. Both studies find that these ‘turn units’ are in fact best thought of as being complex, that is, that they include intonational and pragmatic cues as to where they will end, as well as grammatical ones. Their data show that intonational and pragmatic completion points select from among the many more numerous grammatical completion points to form what we will call, following Ford and Thompson (to appear), ‘Complex Transition Relevance Places’ (CTRPs).
In comparing our three languages for systematic differences in placement of RTs, we found the concept of CTRP to be useful. We counted a Reactive Token as occurring ‘at’ a CTRP if it occurred in the clear (i.e., not in overlap) immediately after the CTRP. In what follows, we will briefly outline our criteria for determining the intonational and grammatical boundaries for identifying CTRPs. 4

4.3.1. Intonational completion

Intonational completion was defined in terms of intonation units. Much previous research has shown that intonation units play a major role in the shape of turn units (cf. Chafe, 1979, 1980, 1987, 1992, 1993, 1994 and Du Bois et al., 1993). We define ‘intonation unit’ as:

"a stretch of speech uttered under a single coherent intonation contour." (Du Bois et al., 1993)

As discussed in Crystal (1969), Cruttenden (1986), Du Bois et al. (1993), Schuetze-Coburn (1992), (to appear), and Schuetze-Coburn et al. (1991), numerous prosodic cues have been identified which are used to determine intonation unit boundaries. The perception of coherence in the pitch pattern is influenced by both the degree and direction of pitch movement on a stressed syllable and by a change in pitch relative to the speaker’s preceding utterance (known as ‘pitch reset’). Timing cues often also play a role in the perception of intonation units, including an acceleration in tempo on initial unstressed syllables, prosodic lengthening of final syllables, and a noticeable pause (0.3 second or greater) between intonation units. The identification of intonation units is thus an auditory, perceptual matter.

In all of our examples, following the transcription system in Du Bois et al. (1993), each line represents one intonation unit. Reactive Tokens generally constitute one intonation unit.

Among the basic intonation unit types, there are two that are characterized as ending in a contour which signals finality, designated by a period or a question mark in the Chafe (1980), (1987), (1994), Du Bois et al. (1993), and Sacks et al. (1974) transcription systems. 5 In contrast, non-final intonation contours are marked in all three systems with commas, and dashes are used to mark intonation contours which break off in mid-utterance.

Each of these four types of intonation unit has its characteristic acoustic realizations in a given language or dialect/variety (Du Bois et al., 1993). For American English and many other languages, the period is realized primarily by a marked fall in pitch at the end of the intonation unit, possibly accompanied by creaky voice. The question mark is realized primarily by a marked high rise in pitch at the end of the

4 Although we have relied on just two of the three criteria used by Oreström (1983) and Ford and Thompson (to appear), namely grammatical and intonational, there is reason to believe that these two criteria capture most of the relevant information. See Oreström (1983) and Ford and Thompson (to appear) for further discussion of incorporating some notion of 'pragmatic completion' into an account of turn transition.

5 These two systems also recognize an exclamatory intonation contour, marked with an exclamation point, which we will treat as a type of final contour for this paper.
intonation unit. The comma, or continuing, intonation unit type is typically realized by a slight rise in pitch at the end of the intonation unit (beginning from a low or mid level), a terminal pitch which remains level, or a terminal pitch which falls slightly but not far enough to be considered final.

Further research will certainly lead to refinements in the specification of the relevant prosodic properties; for the purposes of this project, however, we have taken the intonation unit as a cross-linguistically valid, well-established auditory unit. We divided each transcript into intonation units, either final or non-final, according to the criteria just discussed. An English example can be seen in (9); as noted in the Appendix, final intonation units are indicated by periods or question marks, and non-final intonation units are marked by commas:

(9) *English*
A: and this lady, 
... (H) nobody knows yet why. 
and most of us think, 
she probably fainted. 
... but she fell,  

(AFRICA)

Recall that the two characteristics of a Complex Transition Relevance Place for this project are intonational completion and grammatical completion. What we have shown so far is that for intonational completion, we counted intonation units with a final contour.

4.3.2. Grammatical completion

We judged an utterance to be grammatically complete if, in its sequential context, it could be interpreted as a complete clause, i.e., with an overt or directly recoverable predicate, without considering intonation. In the category of grammatically complete utterances, we also included elliptical clauses and answers to questions.

A grammatical completion point, then, is a point at which the speaker could have stopped and have produced a grammatically complete utterance, though not necessarily one that is accompanied by intonational or interactional completion. As we noted above, our grammatical completion points correspond in practice to what Sacks et al. (1974) called ‘transition-relevance places’. Our definition clarifies the indeterminacy in their characterization, however, since they provide no criteria for identifying one of these points. So a grammatical completion point will sometimes be a word, sometimes a phrase, sometimes a clause, and sometimes a multi-clausal unit, as illustrated in Sacks et al. (1974: 702, fn. 12). An example of an utterance from our data containing a series of grammatical completion points (indicated here by slashes) is the following:

(10) *English*
A: she should have gone/ ho=me/ in= April./ 
.. and she decided to stay/ until Augus=t,/ 
.. (TSK) a=n=d,
the=y went= .. out=/
.. actually,
they just went out/ to Chisera=/
.. to go out/ to the river./ (AFRICA)

In Japanese and Mandarin, we frequently find pragmatic particles following a clause, an NP, and various other grammatical elements. In determining grammatical completion, we coded both the clause without a particle and with the particle as grammatically complete (as indicated with slashes in the examples below).

(11) Japanese
A: ichinichi yasumimashita/ yo./
  one:day was:absent  PRT
  (l) took a day off (OYAMA)

(12) Mandarin
A: .. ni buyao dao Aomen/ qu/ la=/
  2SG NEG to Macao go PRT
  Don't go to Macao. (THAI)

As can be seen from these examples, grammatical completion is evaluated incrementally. That is, a marker of grammatical completion does not necessarily indicate that a complete grammatical unit exists between it and the previous grammatical completion point; rather, as is assumed in the conversation analysis literature, grammatical completion is calculated in terms of its relation with a previous predicate if one is available. Thus, for example, in (10), there is no claim that the adverbial expression in April constitutes an independent unit by itself; rather, it is understood as being a third possible grammatical completion point, the boundary after She should have gone being the first, and that after She should have gone home being the second. By the same token, grammatical incompletion is calculated in terms of a projected upcoming predicate.

4.3.3. Summary
We have outlined two types of completion, i.e., intonational and grammatical, which Oreström (1983) and Ford and Thompson (to appear) have shown to be relevant in characterizing Complex Transition Relevance Places. We will see in section 5 how the three languages of our study differ with respect to the occurrence of Reactive Tokens at CTRPs.

We turn next to our findings.

5. Findings

As noted above, we provide an analysis in terms of the various types of Reactive Tokens as outlined in section 4:
Since there was a substantial amount of variation from one conversation to another within each language, it was not appropriate to pool the data from each language. Accordingly, we are presenting our findings in the form of graphs comparing the eight transcripts in each language.

5.1. Frequency of Reactive Tokens

In order to examine the overall frequency of Reactive Tokens, we consider the ratio of all RTs to all Speaker Changes. That is, we want to ask what proportion of all Speaker Changes serve to support the primary speaker rather than to take the floor for a full turn. Since every RT was counted as a Speaker Change, the ratio of RTs to Speaker Changes provides a frequency measure that is relative to the overall amount of all types of Speaker Change. In fact, the amount of Speaker Change per total number of intonation units was quite similar in the three languages: 28.6% in Japanese, 33.6% in English, and 28.7% in Mandarin. Fig. 1 shows the ratio of RTs to Speaker Changes.

As shown in Fig. 1, with respect to the overall frequency of Reactive Tokens, the three languages of our study show clear differences. These differences can be represented in the form of a hierarchy both in terms of the amount of variation across the transcripts and in terms of the frequency with which RTs are used. This hierarchy is shown in (13):

(13) ENGLISH
    JAPANESE > MANDARIN

Fig. 1 shows that both the range of variation and the ratio of Reactive Tokens to Speaker Change are greater in English and Japanese as compared to Mandarin. The range of variation in English, for example, is from 1.8% to 56.3%, while the range of variation for Japanese is slightly greater, from 3.8% to 66.7%. But Mandarin speakers in our data range from only 1.8% to 26.7% in the number of Speaker

---

6. Our category of 'Resumptive Opener' could be seen as problematic in this regard, since it is followed by a full turn. Nevertheless, we are counting the 'Resumptive Opener' as a Reactive Token and not a full turn, since in itself it does not count as a floor-taking turn, even though it is followed by a TCU which may take the floor.

7. The percentages of Speaker Change in each language if Reactive Tokens are not included is also similar: 17.8% in Japanese, 26.0% in Mandarin, and 21.4% in English.

8. Standard statistical tests assume independent data points. Because our data points are not independent, we have not found it appropriate to run such tests on the results of this project. We therefore refrain from making any technical claims about 'significance', but we do propose that the differences we have found are interesting enough to be presented as findings.
Changes that are RTs, a relatively low percentage and only half the range of variation found in the Japanese and English transcripts. Another way of looking at these findings is to consider the number of transcripts in which at least half of all Speaker Changes are Reactive Tokens: for English there are four such transcripts, and for Japanese there are two. For Mandarin, however, there are no transcripts in which more than half of the Speaker Changes are RTs.

Interestingly, although the range of variation is too great to rely solely on average RT frequency in comparing the three languages, it is worth noting that the averages do support the hierarchy shown in (13), with English RTs comprising an average of 37.3% of the total number of Speaker Changes, Japanese 39.5%, and Mandarin, only 10.0%. Thus in Mandarin, when a new participant speaks, it is more likely to be in order to take the floor than is the case in English and Japanese. Conversely, in English and Japanese, new speakers are more likely to be supporting the primary speaker’s turn than is the case in Mandarin.

In sum, then, our data show a greater use of RTs in Japanese and English conversations than in Mandarin. Japanese and English can be said to exhibit roughly twice the frequency of RT use as Mandarin, and twice the amount of variation. These data allow us to suggest that, based on the frequency of RT use, English and Japanese non-primary speakers seem to play a more active role in supporting the primary speaker, while Mandarin speakers play a less active role; essentially 90% of the Speaker Changes in Mandarin, on the average, serve to take the floor, compared with 60–63% of the Speaker Changes in English and Japanese.

Now we consider the types of Reactive Tokens found in our data.
5.1.1. Distribution of Reactive Token types

Although all RTs are by definition non-floor-taking devices, different languages may utilize different strategies for implementing this support of the primary speaker. Analysis of the different types of RTs allows us to investigate the range of variation and the preferred strategies for supporting the primary speaker in each language.

When we consider the distribution of RT types in the three languages, we find an interesting cross-linguistic similarity: while differences emerge in relative frequencies, for all three languages Backchannels are the most frequent type of RT found in conversation, and Reactive Expressions are the second most frequent. Table 6 presents the average percentages of the different types of Reactive Tokens, showing the similarities and differences among the three languages of our study.

<table>
<thead>
<tr>
<th>Table 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of Reactive Tokens in English, Japanese, and Mandarin</td>
</tr>
<tr>
<td>Language</td>
</tr>
<tr>
<td>Japanese</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>Mandarin</td>
</tr>
</tbody>
</table>

While Backchannels are the preferred form of RT in all three languages, there are striking differences in the relative frequencies of the two or three most frequent types of RTs across the languages.

For Japanese speakers, the obvious favorite RT type is the Backchannel (to be discussed in the next section); 68.3% of all RTs in Japanese are Backchannels. Reactive Expressions and Resumptive Openers are a distant second and third, comprising 17% and 12.5% of all RTs, respectively.

English speakers have a much lower percentage of Backchannels, 37.9%. The next most frequent type of RT, which is more frequent in English than in Japanese, is the Reactive Expression, comprising 34.2% of all RTs. Next most common is the Resumptive Opener, which constitutes a similar percentage of RTs as in Japanese (10.4% vs. 12.5%, respectively).

English speakers have a much lower percentage of Backchannels, 37.9%. The next most frequent type of RT, which is more frequent in English than in Japanese, is the Reactive Expression, comprising 34.2% of all RTs. Next most common is the Resumptive Opener, which constitutes a similar percentage of RTs as in Japanese (10.4% vs. 12.5%, respectively).

The numbers in Mandarin are again too small to allow any strong claims, but we can say that Mandarin speakers are similar to English speakers in their frequency of Backchannels (47.2% of RTs) and Reactive Expressions (31.1%). (Mandarin speakers, like English and Japanese speakers, use Resumptive Openers as their third most frequent type (14.5%), but the number is only 4.)

What we conclude from the data in Table 6, then, is that while relative frequencies of RT types differ from one language to another, there is a similarity across the languages in the types of RTs that are favored by speakers. Backchannels and Reactive Expressions are the most frequent types of RTs in all three languages, suggesting that future research should focus on the interactional functions served by these
favored types. At the same time, Table 6 also points up a striking difference between Japanese, on the one hand, and Mandarin and English, on the other: Japanese more strongly favors Backchannels (68% for Japanese, compared to about 42% for Mandarin and English), while Mandarin and English make greater use of Reactive Expressions than Japanese (about 30% each, compared to 17% for Japanese). The distinction between Backchannels and Reactive Expressions will prove important when comparing our findings on Japanese and English to those of previous researchers.

5.1.2. Backchannel use

As we have seen, overall RT use differs widely among the three languages of our study. Thus, in order to understand the cross-linguistic differences in non-primary speakership, it is important to consider the ratio of Backchannels to all Reactive Tokens. The averages for each language were given above in Table 6, which showed that Japanese speakers greatly favored Backchannels over other RT types, although Backchannels are the favorite RT type in all three languages. Fig. 2 shows the breakdown of Backchannels to Reactive Token ratios by transcript for each language.

As shown in Fig. 2, in Japanese, although there is one transcript with no Backchannels at all, for the speakers in the rest of the transcripts, more than two-thirds of all RTs are Backchannels. In English, the range is between 0% and only 66.7%; in other words, except for the one transcript in each language with no Backchannels, the English maximum percentage of Backchannels is just under the

---

9 Goodwin and Goodwin (1987), discussing the interactional differences between continuers (which are typically 'backchannels' in our terminology) and assessments (which are typically 'reactive expressions' for us), provides a promising place to start.
Japanese minimum. In Mandarin, the range is between 25% and 100%, but the numbers are so small as to make these figures nearly meaningless. (For example, in the transcript in which 100% of the RTs are Backchannels, the number of RTs is one.) It is clear, however, that a much higher percentage of all Reactive Tokens are Backchannels in Japanese than in the other two languages.

Fig. 2 thus suggests the hierarchy shown in (14):

(14) JAPANESE > MANDARIN > ENGLISH

Let us summarize our findings so far. We have chosen a narrower definition of 'Backchannel' than most previous studies, in order to tease out distinctions between the way speakers use all Reactive Tokens and the way they use Backchannel forms. Our data show that, with respect to frequency of RTs, as measured against total Speaker Changes, Japanese and English are comparable, with Mandarin RT use being less frequent. Of all RTs, however, Japanese speakers favor Backchannels to a greater extent than do English and Mandarin speakers.

5.2. Placement of Reactive Tokens

Given that Japanese and English speakers use Reactive Tokens with greater frequency than do Mandarin speakers, our next question is whether there are important differences among the three languages in where speakers choose to place these Reactive Tokens with respect to the primary speaker's turn. When Reactive Tokens are extremely frequent, for example, do they occur while the primary speaker is talking?

Recall that we are accepting the finding from previous research that the basic unit of turn-taking is characterized as ending at a Complex Transition Relevance Place (CTRP), here defined as a point of grammatical and intonational completion. For our analysis of RT location, we will use the CTRP as a way of measuring where speakers place their RTs with respect to the turn of the primary speaker. As with the frequency of Reactive Tokens, the three languages of our study show strong differences in the sequential placement of RTs.

In fact, we found that the extent to which speakers place their Reactive Tokens at CTRPs yields the continuum shown in (15):

(15) MANDARIN > ENGLISH > JAPANESE

That is, the tendency to place RTs at CTRPs is strongest among the Mandarin speakers, weaker among the English speakers, and weakest among the Japanese speakers, as shown in Fig. 3.

Fig. 3 shows that nearly all the RTs, both in individual transcripts, as well as on average, occur at CTRPs in Mandarin, ranging from 44.4% to 100%, though we must continually bear in mind that the numbers of RTs are extremely small. In English (not counting transcript HP, which had only one RT), RTs occurring at CTRPs range from 30% to 66.7% of all RTs, which means that only two of the Mandarin
transcripts have a sufficiently low percentage of RTs at CTRPs to overlap with any transcript in the English range. For Japanese, the transcripts range from 12.1% to 50.0% in percentages of RTs at CTRPs, with no transcript reaching the high end of the English range or even the mid-range of the Mandarin frequency. Thus Fig. 3 shows a clear decrease in preference for placing Reactive Tokens at CTRPs, with Mandarin speakers exhibiting the strongest preference, English the next strongest, and Japanese speakers the least strong preference.\(^\text{10}\) We will see below how this skewing in preferred placement of Reactive Tokens with respect to the interactional structure of the conversation relates to overall strategies in turn management.

But there is another skewing which deserves attention. Degree of preference for placing Reactive Tokens at CTRPs must be understood in terms of the two types of completion that constitute a Complex Transition Relevance Place, namely, intonational completion and grammatical completion. As mentioned above, in all three of our languages, there are many more grammatical completion points in our conversational data than there are intonational completion points. The intonational completion points nearly coincide with CTRPs, but there are many grammatical completion points which do not occur at CTRPs (see Ford and Thompson, to appear and Table 7 below).

So we can ask to what extent Reactive Tokens in the three languages tend to occur at points of grammatical completion, regardless of intonation.

\(^{10}\) Although our percentages are not directly comparable with those of Maynard (1989), they do point in the same direction. Maynard reports that in her data 82.84% of all backchannels in English but only 51.02% in Japanese occur at points of grammatical completion. Maynard’s definition of “grammatical completion” is similar to our definition of CTRP in that it combines intonational (final intonation) and grammatical criteria; however, Maynard does not define grammatical completion incrementally, as we do.
What Fig. 4 shows is that speakers of Mandarin and English strongly tend to place Reactive Tokens at points of grammatical completion, as shown in the hierarchy in (16):

(16) **ENGLISH**

MANDARIN > JAPANESE

That is, fully 88% of Mandarin RTs and 78% of English RTs occur at points of grammatical completion, as compared with only 36.6% of all Japanese RTs. We will explore the implications of this skewing in section 6 below.

Finally, let us consider the implications of the findings on grammatical completion as compared with the results on CTRPs as sites for Reactive Tokens. If we compare Figs. 3 and 4, we see that they are roughly similar for both Mandarin and Japanese. Speakers of both Mandarin and Japanese treat grammatical completion points and CTRPs essentially the same when using Reactive Tokens; that is, the number of Reactive Tokens at CTRPs and at grammatical completion points is roughly the same. But the data for English suggest a substantial preference (roughly twice as great, with a mean of 78.0% compared to 45.1%) for placing Reactive Tokens at grammatical completion points in the primary speaker's turn over placing them at CTRPs. This implies that English speakers are more inclined to wait for the end of a grammatical clause before using a Reactive Token than are speakers of Japanese, and are somewhat less likely to wait for the end of a grammatical clause than are speakers of Mandarin. But, unlike Mandarin speakers, who prefer to place Reactive Tokens at the ends of grammatical clauses that also have final intonation (i.e., at CTRPs), English speakers often use Reactive Tokens at the ends of clauses with non-final intonation contours as well.
While English speakers tend to use Reactive Tokens at the ends of clauses which may or may not be intonationally complete, and Mandarin speakers tend to use them at points of both grammatical and intonational completion, it is clear that Japanese speakers must be using Reactive Tokens at points lacking either grammatical or intonational completion. In Japanese, since only 30.8% of all Reactive Tokens occur at CTRPs and only 36.6% at Grammatical Completion Points, the question arises: where are the rest of the Reactive Tokens located? The answer is that more than half of all Japanese RTs occur in the following non-final locations: 19.6% in the middle of intonation units, and 33.3% at the ends of intonation units which lack grammatical completion and which have non-final intonation contours (marked with commas in our transcripts). Thus Japanese RTs are much more likely to occur 'midstream' during the primary speaker's talk, while s/he is still in the process of constructing a grammatical clause.

The following example, in which two co-workers are complaining about their jobs, illustrates these two types of non-final Reactive Tokens in Japanese.

(17) Japanese

T: wareware no ne=, we GEN PRT
    we
    our

H: un.
    uh-huh

T: sofuto no ne=, software GEN PRT
    on software

H: un.
    uh-huh

T: ... shigoto ni taisuru hyoo[ka] ga, work LOC towards recognition SUB
    recognition of the work
    [n].
    uh-huh

T: ano hito ne=, that person PRT
    that person
    .. shite nai.
    do not
    doesn't
That person doesn't recognize our software work.

At the beginning of (17) we see a very typical pattern of RTs after non-final intonation units in the Japanese data. Following T's first two intonation units, each of which consists of a noun, casemaker and pragmatic particle, H produces a backchannel. Then, in the less frequent 'midstream' pattern, H backchannels during T's third intonation unit, in the middle of the noun hyooka 'evaluation'. Clearly, H has established a steady rhythm of backchanneling at or near the end of each intonation unit as T produces it; the fact that the final backchannel precedes the end of the intonation unit may reflect the predictability of the word hyooka 'recognition' in this context, especially after the first syllable has been produced, as well as its crucial role in completing the noun phrase 'evaluation of our software work'. But since all three backchannels are produced before T completes a single clause, H is not reacting to a full proposition per se, even though he may be able to anticipate the direction of T's thoughts.

In (18) we see a 'midstream' Reactive Token that involves a rather extensive overlap. H and T have been talking about T's father (H's former teacher), who unexpectedly went off to sleep somewhere during H's visit; H has asked who slept where.

(18) *Japanese*

H: chanto=, properly

T: un. uh-huh

H: a, oh

okita [1 kara heki 1] [2 datta 2] n da.
got:up because all:right was NOM COP

he got up so it was all right.

T: [1 okite kita 1]

got:up came (he) got up

[2 un. 2]

uh-huh

---

12 This noun is repeated after the next two intonation units, probably because current speaker T has used the wrong casemaker, the subject particle ga, instead of the object particle o.
After H begins the answer to T’s question, T immediately produces a partial repetition in overlap with her, and then continues to overlap with a backchannel. This type of rather long overlapping Reactive Token has also been documented in Japanese by Hayashi (1988), who found more than twice as much simultaneous talk among Japanese as among American conversationalists. Although rather common in Japanese, in English and Mandarin the production of RTs after non-final intonation units or ‘midstream’ during intonation units is extremely rare.

In this section, then, we have seen that speakers of the three languages of our study differ in the extent to which they place RTs at points of grammatical and intonational completion, i.e., what we are calling CTRPs. We found that the tendency to place RTs at CTRPs is strongest among the Mandarin speakers, weaker among the English speakers, and weakest among the Japanese speakers, as shown in Fig. 3 above.

We also found that speakers of Mandarin and Japanese treat grammatical completion points and CTRPs essentially the same when using Reactive Tokens, while the data for English suggest a substantial preference for placing Reactive Tokens at grammatical completion points in the primary speaker’s turn over placing them at CTRPs, as shown in Fig. 4 above.

In section 6 we will discuss our findings on the placement of Reactive Tokens in terms of overall strategies in turn management. But there is one final issue to be resolved before turning to the implications of our findings.

5.3. Behavior of the non-primary speaker at boundary points

Thus far we have been analyzing the treatment of various boundary points, such as CTRPs and points of grammatical completion, in terms of the percentage of Reactive Tokens at each point. Before we consider the implications of our findings on RTs for interactional strategies, it is important to have an overview of how the different boundary points are handled in conversation within each language. In Table 7, therefore, we summarize the behavior of the non-primary speaker at four different boundary points in our data: CTRPs (points of grammatical and intonational completion), grammatical completion points (the ends of complete grammatical clauses), intonational completion points (points of final intonation), and non-final intonation units (i.e., points marked by commas in our transcription system). We consider three possible moves of the non-primary speaker: (1) produce a Reactive Token, (2) take a full turn, (3) no verbal reaction (indicated by 0 in Table 7).

As the figures in the last row of Table 7 show, the most frequent reaction of the non-primary Speaker at each of these boundary points is simply to continue listening, with no verbal response. There is a striking difference between the points of non-final intonation, however, and the other boundaries: as might be expected, much lower rates of Reactive Tokens and Speaker Change occur at the ends of non-final intonation units than at the boundaries involving grammatical and/or intonational completion. Only 8% or fewer non-final intonation units elicit either a Reactive Token or Speaker Change, with Mandarin conversationalists being the most likely to continue listening in silence (94% of the time).
Table 7
Non-primary Speaker Behavior at Boundary Points

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J</td>
<td>E</td>
<td>M</td>
<td>J</td>
</tr>
<tr>
<td>RT</td>
<td>0.08</td>
<td>0.14</td>
<td>0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>Turn</td>
<td>0.29</td>
<td>0.39</td>
<td>0.39</td>
<td>0.28</td>
</tr>
<tr>
<td>0</td>
<td>0.63</td>
<td>0.48</td>
<td>0.56</td>
<td>0.63</td>
</tr>
</tbody>
</table>

At the three major completion points in Table 7, the non-primary speaker in all three languages is more likely to take a full turn than to give a Reactive Token. At CTRPs and points of intonational completion, English and Mandarin non-primary speakers take a turn almost 40% of the time, Japanese non-primary speakers almost 30% of the time. Reactive tokens are provided about 5–15% of the time at these boundaries, with English showing the highest and Mandarin the lowest rates of RTs. Points of grammatical completion show the same pattern, but with somewhat lower rates of Speaker Change and Reactive Tokens.

In sum, what the figures in Table 7 allow us to do is to compare the use of RTs in the three languages against a different measure from that used for Figs. 3 and 4. Figs. 3 and 4 considered RT placement in terms of percentages of all RTs, while Table 7 considers RT placement in terms of all boundary points. These data taken together indicate, then, that although the rates of RT use at various boundary points may be high, as shown in Figs. 3 and 4, when we consider RT use in terms of the four types of boundary points, we can see that the ratio of RTs per boundary point may be very low. For example, while 80% of Mandarin speakers' RTs occur at CTRPs (Fig. 3), this is by no means the same as saying that 80% of CTRPs have RTs uttered after them. Indeed, only 5% of Mandarin CTRPs are followed by an RT (Table 7). And although 33.3% of Japanese RTs occur at the ends of non-final intonation units (see p. 375), this constitutes only 8% of the total number of these units. Table 7 thus clarifies the relationship between RT use and boundaries in the primary speaker’s talk.

5.4. Summary

We have seen that, in terms of both frequency and placement of Reactive Tokens, our Japanese, Mandarin, and English data show clear and interesting differences. First, Japanese and English speakers use Reactive Tokens more than three times as frequently as Mandarin speakers (Fig. 1), which we represented by the hierarchy in (13):

(13) ENGLISH
     JAPANESE > MANDARIN

Second, in terms of ratio of Backchannel responses to total RTs, Japanese (68.3%) outranks Mandarin (47.2%) and English (37.9%) (Fig. 2), as shown in (14):
Third, Mandarin speakers place a higher percentage of their Reactive Tokens at Complex Transition Relevance Places than English speakers, who in turn favor CTRPs more than Japanese speakers (Fig. 3), as shown in the hierarchy in (15):

(15) MANDARIN > ENGLISH > JAPANESE

Finally, Mandarin and English speakers place higher percentages of RTs at points of grammatical completion than do Japanese speakers (Fig. 4), producing the hierarchy in (16):

(16) ENGLISH > JAPANESE
MANDARIN

Summary profiles of these findings by language are given in Table 8.

Table 8
Summary profiles of the non-primary speaker in Japanese, Mandarin and English

<table>
<thead>
<tr>
<th>Japanese</th>
<th>Mandarin</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Reactive Token ratio</td>
<td>Low Reactive Token ratio</td>
<td>High Reactive Token ratio</td>
</tr>
<tr>
<td>High Backchannel ratio</td>
<td>Low Backchannel ratio</td>
<td>Low Backchannel ratio</td>
</tr>
<tr>
<td>Low RT at CTRP</td>
<td>High RT at CTRP</td>
<td>Moderate RT at CTRP</td>
</tr>
<tr>
<td>Low RT at gram. compl.</td>
<td>High RT at gram. compl.</td>
<td>High RT at gram. compl.</td>
</tr>
</tbody>
</table>

In the next section, we interpret these results in terms of differences in overall turn-management strategies among the three languages.

6. Implications

Our findings with respect to Reactive Token use in these three languages provide a new type of evidence for the observation that what it means to be a 'polite' or 'co-operative' conversational partner (Brown and Levinson, 1978; Grice, 1975) is a culture-specific matter, as suggested in Keenan (1976), Rosaldo (1982), and Wierzbicka (1985), among others.

We are well aware that there must be other factors which influence RT use which we have not examined. Our data consist of conversations among friends of the same social status, so they do not allow us to generalize about the role of social status in RT use. Similarly, gender may well play a role; there are no clear indications of this in our data, but we did not control for gender in our research design, so our data simply leave us agnostic on this point. Finally, we readily acknowledge that there will almost certainly be individual differences among speakers within a given language.
arising from the content and setting of the conversation. Investigating the role of these various factors is a task for continued research.

The findings which we have summarized on the use of Reactive Tokens suggest that they are part of a distinct set of turn-management strategies for each language. That is, the frequency, types, and placement of Reactive Tokens are part of an interactional system which competent language users know and which gives rise to clear cultural expectations about what speakers and listeners are doing in ordinary talk, similar to the conventionalized styles documented for various cultural subsets of American speakers in Tannen (1981a,b,c, 1982, 1984, 1987). In what follows we will try to outline what our study shows about interactive strategies in each of the three languages we have investigated.

6.1. Japanese

There is a relatively substantial literature suggesting that interactional and affective considerations play a larger role in Japanese conversations than they do in English. In Japanese language teaching, how to teach the appropriate forms, functions, and use of Reactive Tokens is a very important issue (Horiguchi, 1988; Matsuda, 1988; Mizutani, 1983, 1984, 1988). The use of *aizuchi* 'backchannels' is a matter of everyday discussion among Japanese people; it is common to comment on other people's over- or under-use of *aizuchi*. Furthermore, there are many anecdotal reports by native speakers of Japanese and English suggesting that a higher rate of Reactive Token use tends to characterize Japanese conversation.

Our findings are consistent with prior research in that Japanese shows a high frequency of Reactive Tokens, but we did not find that this frequency was significantly higher than in English. In contrast, Maynard (1989), based on 3-minute segments of conversation among 20 Japanese and 20 American pairs, found a total of 871 Japanese 'backchannels', occurring roughly one per every 2.42 Pause-bounded Phrasal Unit, compared with only 428 American 'backchannels'. In our data, which are based on fewer, though somewhat longer stretches of conversation, we found roughly equivalent frequencies of Reactive Tokens in Japanese and English (39.5% and 37.3% of all Speaker Changes, and 10.8% vs. 10.9% of all Intonation units, respectively). When we consider only Backchannels, since a higher percentage of Japanese Reactive Tokens are Backchannels, we find a greater difference: in Japanese, 29.9% of all Speaker Changes were Backchannels, while only 15.9% of all Speaker Changes were Backchannels in English.

Another important difference in our Japanese vs. English data is the location of Reactive Tokens. As noted above, Americans place 45.1% of their RTs at CTRPs and 78.0% at Grammatical Completion points, in contrast to Japanese speakers, who instead place more than half of all their RTs in the middle of the primary speaker's IU or at the end of non-clausal IUs. Thus our Japanese speakers are much more likely to give their Reactive Tokens while the primary speaker is 'in progress' rather than waiting for a completion point. Regardless of relative frequency, this unexpected placement of RTs could lead Americans to feel that their Japanese interlocutors are using Reactive Tokens much more than anticipated.
In analyzing the implications of such differences for interaction, it is important to keep in mind that culture-specific interpretations of the same phenomenon often vary radically. Our data suggest that Americans expect to hear Reactive Tokens at Grammatical Completion points; from a semantic/pragmatic point of view, these are presumably points at which a full proposition has been uttered. If the listener does not wait to hear a complete clause before giving a Reactive Token, this may sound dismissive ("Yeah, yeah, I already know what you mean"), and may be disruptive to the speaker, who is not accustomed to processing a reaction from the listener while formulating a clause. In this light it makes sense that if RTs are to be produced while the speaker is in progress, as in Japanese, their most common form should be the Backchannel, which presumably places the least burden on the primary speaker, since it does not require any syntactic/semantic analysis, and is taken as a continuer (Schegloff, 1982).

On the other hand, from the Japanese perspective, Reactive Tokens that occur while the speaker is in progress rather than finished with a grammatical or intonational unit may be ideally suited for providing emotional support; to the extent that RTs do not occur at the ends of propositions, they are unlikely to be interpreted as providing support for the speaker's point, which has not yet been made. Mizutani (1984, 1988) and Maynard (1989) suggest that the Japanese turn-management strategy could be characterized as follows: speakers use a highly conventionalized affect-laden interactional style, with frequent inviting and accepting of involvement on the part of all participants. Thus, Japanese speakers use RTs as one way of showing interactional support (see, e.g., Horiguchi, 1988; Mizutani, 1983). Matsuda (1988) points out that this characterization may be related to a Japanese cultural concern for harmony and co-operation.

Our data provide support for this description, and suggest that Reactive Tokens which occur frequently and are distributed throughout another speaker's turns and clauses rather than at possible completion points may constitute an especially appropriate means of providing and receiving interactional support in conversation. The non-primary speaker is expected to show concern for the primary speaker's sense of security in holding the floor; an empathetic conversational partner provides this sense of security by giving RTs during the primary speaker's turn.

6.2. Mandarin

In sharp contrast to the research tradition in Japanese, there has been no systematic work done on Reactive Tokens in Mandarin or any other Chinese language. This by itself is an interesting fact, suggesting that Reactive Token use is not perceived as a predominant feature of Chinese language use. The literature discussed above (Mizuno, 1988, and Liu, 1987) does suggest that Chinese speakers use aizuchi ('backchannels') less frequently than do Japanese speakers. Our findings confirm these results for Backchannels and Reactive Tokens in general, and indicate a quite different turn-management strategy than for Japanese (or English). First, the use of Reactive Tokens is strikingly low in comparison with Japanese and English. Second, when the Mandarin speaker does react, s/he tends to do it at a Complex Transition
Relevance Place and not in the middle of the primary speaker's clause. And when a Reactive Token is used, it is much more likely to be lexically contentful than are Japanese Reactive Tokens; this places no special burden on the primary speaker, who has presumably completed his/her clause, and is free to perform any syntactic/semantic analysis required.

Based on informal observations from several bilingual Mandarin speakers, our speculation is that Mandarin interactional style favors conversational participants not infringing on the other's 'turn space'. RT use, particularly without waiting for a transition point, is seen as presumptuous, intrusive, and even rude or impolite. An interesting parallel can be found in the work of Philips (1983: 52–67), who reports that in the Warm Springs community, Native American listeners indicate their attention to the speaker by very subtle movements of the muscles around the eyes, and use fewer backchannels than Anglo Americans. This listener behavior is part of a 'non-coercive cultural orientation' that places high value on personal autonomy and avoids putting oneself above others; listeners avoid behaviors that might imply an attempt to control the speaker or to indicate that they are listening more attentively than others. Thus avoidance of backchanneling is interpreted as reflecting an appropriate stance of non-interference toward the speaker. For Japanese listeners, failing to provide RTs for the primary speaker may be interpreted as uncooperative and lacking in empathy - a failure to care for the speaker's needs - while in Mandarin, eschewing RTs shows an appropriate respect for the primary speakers' right to formulate and produce their talk undisturbed. This interpretation of our results shows an intriguing parallel with notions discussed in terms of politeness, e.g. R. Lakoff's (1973) 'rules of politeness' which specify an opposition between non-imposition/freedom ('Don't impose' and 'Give options') and camaraderie ('Make A feel good - be friendly') and Brown and Levinson's (1987: 67) distinction between 'negative face' (the wish for one's actions to be unimpeded by others) and 'positive face' (the wish for one's wants to be desirable, i.e., the wish to be appreciated). These speculations must await further study; for now we merely note that the use of Reactive Tokens in conversation may prove to be highly consistent with such culture-specific interactional phenomena as politeness strategies.

6.3. English

In very broad terms, English could be said to occupy a position between Japanese and Mandarin with respect to Reactive Token use. The relatively high frequency of Reactive Tokens suggests a strongly interactional style with numerous reactions on the part of the non-primary speaker. Reactive Tokens are frequent and often occur within another speaker's turn. But the American interactive style differs from the Japanese in that speakers do not usually provide Reactive Tokens until a point of grammatical completion; it differs from the Mandarin in that Americans do not necessarily wait until a CTRP has been reached before they react. Reactive Tokens in our English data are often contentful, requiring a certain minimal amount of linguistic processing, but the fact that they are usually produced at points of grammatical completion minimizes any potential difficulty in processing them.
Based again on purely anecdotal accounts, we note that Americans are indeed caught in the middle with respect to Japanese and Mandarin Reactive Tokens: they tend to find Japanese RTs disruptive and even annoying, but the Mandarin paucity of RTs somewhat unnerving, leaving them wondering what the listener is thinking. Americans are not accustomed to Japanese-style ‘midstream’ affective support, but miss the punctuating of a large number of their propositions, even non-final ones, with a Reactive Token from the listener. Thus American non-primary speakers are more actively involved than their Mandarin counterparts, but are nevertheless expected to refrain from infringing on the primary speaker’s on-going task of formulating propositions. In terms of interactive style, then, Americans may stand at a point somewhere between the high rapport of Japanese and the respectful deference of Mandarin interactants.

7. Conclusions

In this study we have shown how a close examination of one interactional device, the Reactive Token, can shed light on possible cross-linguistic differences in communicative strategies. A major finding is that differences from one conversation to another in our data could be as great as differences across languages. Yet when these differences are taken into account, we can still see patterns of usage which characterize each language as opposed to the others.

We see our study as a first step towards a more systematic understanding of the extent of possible variation from one language to another in the way people carry on everyday conversations. One obvious question for further research relates to the factors underlying the variation in RT use across conversations within one language. For example, what is it about the English ‘HYPO’ or the Japanese ‘OYAMA’ conversations that leads to such a relatively low use of RTs? Since all the conversations were among friends, relative status is not a likely factor; we speculate that a variety of factors might be found to relate to differential use of RTs, such as content, number of speakers, or ethnic or regional ‘style’, as discussed in Tannen (1981a,b,c, 1982, 1984, 1987), but studies attempting to determine the role of such factors have yet to be done.

Since our analysis has been primarily quantitative, we have not carried out the type of detailed case-by-case analysis of RTs in a given language that would establish the sequential and organizational properties that almost certainly affect their frequency and distribution, as suggested and demonstrated in, e.g., Goodwin (1986), Goodwin and Goodwin (1992b), Jefferson (1984), and Schegloff (1982). We recognize, of course, that such in-depth analyses are critical for an understanding of the way RTs work in interaction and hope that our current project will provide partial guidance for future research in this area. Thus starting from our results, we might ask what sequential properties of conversations in these three languages could help account for the more global differences suggested by our findings. For example, do certain types of conversational sequences occur with different frequencies in different cultures? Does conversation involving similar topics and participants exhibit dif-
ferent sequential organization in different cultures? We hope that our quantitative approach will soon be complemented by more qualitative case-by-case analyses of Reactive Tokens across different cultures.

We also anticipate further research to determine the range of functions that RTs serve from one language to another, including the types of functions that have been discussed as ‘assessments’ (C. Goodwin, 1986; M. Goodwin, 1980; Goodwin and Goodwin, 1987, 1992a,b), and ‘continuers’ (Schegloff, 1982). And we would welcome further research showing how such findings could be incorporated into our understanding of the ways in which languages are learned and taught.

What we hope to have shown is that such cross-linguistic comparisons are feasible and can begin to open new lines of inquiry into conversational strategies. While our cross-cultural interpretations of the data await confirmation from further research, it is clear that cross-linguistic analysis of Reactive Tokens can make a valuable contribution to our understanding of conversational interaction.

Appendix: Symbols for discourse transcription from Du Bois (1991) and Du Bois et al. (1993)

\begin{itemize}
  \item \textbf{Units}
  \begin{itemize}
    \item Intonation unit \{carriage return\}
    \item Truncated intonation unit \-\-
    \item Word \textquoteleft space\textquoteright
    \item Truncated word \-
  \end{itemize}

  \item \textbf{Speakers}
  \begin{itemize}
    \item Speaker identity/turn start : \\
    \item Speech overlap [ ]*
  \end{itemize}

  \item \textbf{Transitional continuity}
  \begin{itemize}
    \item Final .
    \item Continuing ,
    \item Appeal ?
  \end{itemize}

  \item \textbf{Lengthening}
  \begin{itemize}
    \item =
  \end{itemize}

  \item \textbf{Pause}
  \begin{itemize}
    \item Long ...
    \item Short ..
  \end{itemize}

  \item \textbf{Vocal noises}
  \begin{itemize}
    \item Vocal noises ( )
    \item Alveolar click (TSK)
    \item Inhalation (H)
  \end{itemize}

  \item \textbf{Quality}
  \begin{itemize}
    \item Quality <Y Y>
  \end{itemize}
\end{itemize}
**Transcriber’s perspective**

Uncertain hearing

* Certain brackets are indexed with numbers to clarify which speech overlaps with which.

**References**


