Volume 31 • No. 1 May, 2009

Japan Association for Language Teaching

全国語学教育学会

jalt journal

Articles

Reviews

ISSN 0287-2420

¥950

JALT Journal

Volume 31 • No. 1
May, 2009

Contents

May 2009
Volume 31 • No. 1

Japan Association for Language Teaching

A Nonprofit Organization

The Japan Association for Language Teaching (JALT) is a nonprofit professional organization dedicated to the improvement of language teaching and learning in Japan. It provides a forum for the exchange of new ideas and techniques and a means of keeping informed about developments in the rapidly changing field of second and foreign language education. Established in 1976, JALT serves an international membership of approximately 3,000 language teachers. There are 34 JALT chapters, all in Japan, along with 16 special interest groups (SIGs), two forming chapters, and four forming SIGs. JALT is one of the founders of PAC (Pan-Asian Consortium), which is an association of language teacher organizations in Pacific Asia. PAC holds regional conferences and exchanges information among its member organizations. JALT is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and is a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes *JALT Journal,* a research journal; *The Language Teacher*, a monthly periodical containing articles, teaching activities, reviews, and announcements about professional concerns; and *JALT International Conference Proceedings*.

The JALT International Conference on Language Teaching and Learning and Educational Materials Exposition attracts some 2,000 participants annually and offers over 600 papers, workshops, colloquia, and poster sessions. Each JALT chapter holds local meetings and JALT's SIGs provide information on specific areas of interest. JALT also sponsors special events such as workshops and conferences on specific themes, and awards annual grants for research projects related to language teaching and learning. Membership is open to those interested in language education and includes automatic assignment to the nearest chapter or the chapter you prefer to join, copies of JALT publications, and reduced admission to JALT-sponsored events. JALT members can join as many SIGs as they wish for an annual fee of ¥1,500 per SIG. For information, contact the JALT Central Office or visit the JALT website at <http://www.jalt.org>.

JALT National Officers, 2009

President: . . . . . Caroline Lloyd
Vice President: . . . . . Cynthia Keith
Auditor: . . . . . Tadashi Ishida
Director of Treasury: . . . . . Kevin Cleary
Director of Records: . . . . . Aleda Krause
Director of Programs: . . . . . Philip McCasland
Director of Membership: . . . . . Nathan Furuya
Director of Public Relations: . . . . . Emi Itoi

Chapters

Akita, Chiba, East Shikoku, Fukui, Fukuoka, Gifu, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagoshima, Kitakyushu, Kobe, Kyoto, Matsuyama, Miyazaki, Nagasaki, Nagoya, Nara, Niigata (forming), Okayama, Okinawa, Omiya, Osaka, Sendai, Shinshu, Shizuoka, Tokyo, Toyohashi, Wakayama (forming), West Tokyo, Yamagata, Yokohama.

Special Interest Groups

Bilingualism; College and University Educators; Computer-Assisted Language Learning; Extensive Reading (forming); Gender Awareness in Language Education; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Lifelong Language Learning; Materials Writers; Other-Language Educators; Pragmatics; Professionalism, Administration, and Leadership in Education; Study Abroad (forming); Teacher Education; Teachers Helping Teachers (forming); Teaching Children; Testing and Evaluation.

JALT Central Office

Urban Edge Building, 5F 1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan

Tel.: 03-3837-1630; Fax: 03-3837-1631; E-mail: jco@jalt.org;

Website: <www.jalt.org>

In this Issue

Articles

Conversation Analysis (CA) has continued to gain traction within Applied Linguistics as a micro-analytic approach to understanding human interaction. This issue opens with a contribution in this area from **Tim Greer**, **Vivian Bussinguer S. Andrade**, **Jeff Butterfield**, and **Agnes Mischinger**. The study examines the interactional practice of repetition as recipiency in the L2 context. In our second article, **Kyoko Miyazato** explores the politics of team teaching (TT). From a naturalistic case study she draws conclusions which contribute to our understanding of the relationship between Assistant English Teachers (AETs) and Japanese Teachers of English (JTEs). Our third article, by **Keith Ford**, uses semi-structured interviews to examine the use of L1 in the L2 classroom. Our final English language article, by **Takaaki Kumazawa**, reports a study using generalizability theory (G theory) and a decision study (D study) to analyze a criterion-referenced vocabulary test. Methodology which proceeds beyond the limitations of classical test theory in its ability to analyze sources of error is illustrated in practice. In our Japanese language contribution, **Yuko Nakahama** examines the influence of task complexity, learners' L1 background and proficiency levels on the introduction and maintenance of referential topics in L2 Japanese oral narratives.

Reviews

In the first of our five book reviews, **Tim Greer** reports on a book offering a critical overview of bilingual people and code switching. In the second, **Ian MacLean** reports on a book which addresses academic and instructional issues regarding idioms. Our third review, by **Scott Gardner**, considers a volume on English language learning materials and how developers have responded to calls for more student-oriented materials. Next, **Patrick Foss** reports on a book which offers a positive take on English-based loanwords in the Japanese lexicon and their place in English language education. Finally, **Andre A. Parsons** reports on a volume examining aspects of professional talk encountered by teachers at different points in their careers.

From the Editor

This issue of *JALT Journal* sees **Bill Perry** joining us as the new Reviews Editor. As I mentioned in the previous issue, **Yuriko Kite** has retired from the position and we thank her for the contribution to the journal. I also welcome Yuriko to the Editorial Advisory Board.

As always, I would like to thank the members of the Editorial Advisory Board as well as our additional readers, and of course the proofreaders, for continued support. *JALT Journal* represents the sum of all your effort.

Errata

We would like to apologize for an error in our previous issue (November, 2008). In **Mark Rebuck**'s review of the book *Shogakusei ni eigo wo oshierutowa? Ajia to nihon no kyouiku genba kara* [*What is the Meaning of Teaching English to Elementary School Children?*] by **Kawahara Toshiaki**, the Japanese word "Kakusa" was printed incorrectly as "Kakusai."

Articles

Receipt Through Repetition

Tim Greer
Vivian Bussinguer S. Andrade
Jeff Butterfield
Agnes Mischinger
Kobe University

One of the ways that people show they are listening is by repeating part of what the prior speaker just said. This practice allows listeners to establish recipiency in a way that is specific, providing the speaker with moment-by-moment feedback on the recipient's understanding as well as giving a "go-ahead" signal. This paper uses Conversation Analysis (CA) to explore the interactional practice of repetition as recipiency. The data are taken from video recordings of L2 users of English in paired and small group discussions. The analysis provides suggestions for how this practice can help encourage language learning in conversation settings.

相手が言ったことの一部を繰り返すことは相手の言っていることを聴いているということを示す方法の一つである。相手の発話を繰り返すことにより聞き手は自分が相手の言うことを理解しており、それを承認している、さらに続けられたしということを刻々とフィードバックしていることになるのである。このようにして聞き手は相手の言うことを受容（recipiency）することになるわけである。本研究は会話分析（Conversational Analysis）を分析方法として使用し、繰り返しによる受容（receipt through repetition）の構造を明らかにする。英語学習者がグループ活動を行っている際の会話をビデオに録画したものをデータとして使用した。会話での繰り返しが話し手と聞き手相互の共通理解を生みだし、言語学習をサポートする経過が明らかにされた。

Keywords: recipiency, repetition, conversation analysis, interaction, novice talk, L2 pragmatics

As language teachers, we are constantly observing students speaking in their second language. Intuitively we know when they are succeeding and when they are having trouble, or when they seem to be doing something in a way that is different from the way that a 'native speaker' might do it. However, it is often difficult to spell out just what such things are.

With its participant-centered focus on naturally occurring interaction, Conversation Analysis (CA) is one methodological approach that is currently receiving increasing interest among applied linguists as a means of documenting how novice speakers accomplish various social actions in their second language (most notably, Firth and Wagner, 1997; Gardner & Wagner, 2004; Markee & Kasper, 2004; Richards and Seedhouse, 2005; Schegloff, Koshik, Jacoby, & Olsher, 2002; Seedhouse, 2005; Wagner and Firth, 2007). Just as natural sciences like astronomy or geology are built on empirical descriptions of natural phenomena, so too does CA aim to provide a detailed descriptive account of an ordinary observable occurrence: interaction. CA generally gathers its evidence from how speakers act and react to the turn-by-turn sequential development of mundane and institutional conversations (Schegloff, 1996a). Using video and audio recordings, CA researchers develop a case by gathering collections of similar interactional phenomena and describing them from the micro-socio-perspective of the participants themselves.

In the current study we examine video-recorded interactional data between Japanese learners of English during speaking tests. We initially noticed that the participants often repeated words, either within their own turn or as a re-doing of some element of the prior speaker's turn. By paying careful attention to the sequential contexts in which these repetitions were employed, we found that, despite their limited ability, the novice learners used next-turn repetition in much the same ways as relative experts, such as English native speakers:[1] to *initiate repair*,[2] to agree, and to claim comprehension of the topic at hand. The last of these is the focus of the current paper: receipt through repetition.

CA Research into Receipt Markers and Repetition

We will begin by reviewing previous socio-interactional research into receipt and repetition. By *receipt* we are referring to those minimal turns at talk which demonstrate that a person is listening. Such 'reactive tokens' do not stop the primary speaker from talking, and do not in themselves claim the floor (Clancy, Thompson, Suzuki & Tao. 1996, p. 356). Tokens that

most regularly accomplish this sort of action in English include *yeah*, *mm*, *uhuh* and *I see*, and in Japanese *un*, *hai*, and *ohn*. They have been identified in the literature under a variety of terms, including *accompaniment signals* (Kendon, 1967), *backchannels* (Yngve, 1970), and *aizuchi* (Maynard, 1989). While much of the existing research, including CA studies (Jefferson, 1985, 1993; Gardner 1997, 2001), focuses on L1 interaction, some preliminary research has been conducted on receipt tokens among second language learners. Gardner (1998), for example, looked at how learners vocalize their understanding by using the acknowledgement token 'mm.'

While such tokens are by far the most common way to do receipt, they are not the only means speakers have of demonstrating understanding. In fact there are a number of related—but slightly different—practices that interactants use to do receipt (Schegloff, 1982). Utterances like *wow* or *good* for example, offer some assessment of the prior turn (Goodwin, 1986); or the *change-of-state token* 'oh' makes a claim that the recipient has undergone an epistemological transition—from 'not-knowing' to 'now-knowing' (Heritage, 1984; Schegloff, 2007). The fact that these forms of receipt change the way the primary-speaker produces the remainder of the ongoing talk is evidence to suggest that not all receipt tokens accomplish listenership in the same way. While fascinating, these recipiency practices will remain largely beyond the scope of the present study.

The receipt practice in which we are interested here involves repeating some element of the primary-speaker's turn. Schegloff (1996b) notes that such receipts are produced with downward intonation, and are often followed by agreement, acknowledgement, or confirmation tokens, such as in excerpt 1.

Excerpt 1: Kanagawa (Receipt-through-repetition)

```
01 C:      where where are you,
02 D:      mm.
03 C:      from.
04            (0.8)
05 D:      ahm:m(0.6)>kanagawa.<
06 C: →    >kanagawa.<  oh[n
07 D:                     [ah=
08 C:      =mmm
```

In line 6 of this excerpt, C provides receipt of D's prior turn by repeating the key element with falling intonation, and following it immediately with an acknowledgement token in Japanese. D then treats the repeat as response-worthy (Schegloff, 1996b) by doing confirmation in line 7. Accomplishing listenership can happen this way in most languages, but Clancy et al. (1996) found that repetitions were used as reactive tokens almost twice as often in Japanese as in English, making it probable that Japanese learners of English, like those in the present study, will tend to over-rely on this interactional practice.

The practice of other-repetition as receipt has been studied from a CA approach by Svennevig (2004), who found that the recipient could use next-turn repetition to make public a variety of inner cognitive states, including claims to hearing and understanding, or to express an emotional stance in regard to the prior speaker's turn. It is this focus on outwardly observable, real-time claims to otherwise inaccessible 'in-the-head' states that makes CA invaluable for investigating language learning as a socioculturally accomplished process (Wagner and Firth, 2007). Svennevig's data was collected from institutional talk between expert and novice speakers of Norwegian, and will inform the current study, which focuses only on novice speakers of English.

However, just as not all receipts are done through repetition, neither do all repetitions accomplish receipt. Next-turn other-repetitions are regularly used with upward intonation to seek confirmation or initiate repair, such as in the turns marked with an arrow in the following extracts.

Excerpt 2. Initiating a confirmation check

```
01 B:    >ah but< (0.5) this spring (0.5)
02       I went- I go to: (.) um Hong Kong
03       (0.3)
04 C:→   Hong Kong?
05       (0.2)
06 A:    [oh ]
07 B:    [yes] with my friends.
```

Excerpt 3. Next-turn repair initiation through repetition

```
01 A:    um:: a:nd umm (0.7) in Otaru,
02       they er there are many (0.4) slope.
```

```
03          (0.4)
04 D:→      slope?
05          (0.8)
06 D:       what slope?
07          (0.7)
08 A:       >s::aka.   saka.<
             slope     slope
```

Notice that when some element of the prior turn is repeated with rising intonation, as in these examples, the next-speaker hears it as some sort of sequence-initiating action and responds to it, whereas next-turn repetition like that in excerpt 1 closes an action sequence. In CA terms, the former is the first pair part of an adjacency pair (e.g., Question/Answer), while the latter is a sequence-closing third, since it acknowledges the second pair part of a just-prior adjacency pair (see Schegloff, 2007). In short, repetition delivered with upward intonation is doing the opposite of receipt: the listener is making a claim that the prior information is unknown, unrecognized, untrue, or in some other way problematic.

Repetition can also be used to accomplish agreement (Pomerantz, 1984):

Excerpt 4. Repetition as agreement

```
01 A:       why didju apply to, Hokudai.
02          (1.3)
03 B:       ah:: (0.5) I love (0.2) this, (0.5) la:rge-i,=
04 A:       =ah:[:
05 C:           [campus.
06 A:→      oh (.) ah very la(h)rge(h).
07 B:       .hh heh [aha ((nodding)
08 D:               [°u[n.°]= ((nodding))
09 A:                  [un.] ((nodding))
10 C:       °yes:::.°
```

Here, in line 3 speaker B produces an assessment (‘large’), which receives immediate acknowledgement from A in next-turn and then, in line 6, an upgraded repetition of the assessment (‘very large’) and multiple agreement

tokens and nods from the other recipients during the interaction that follows.

Another location where next-turn repetition was regularly found in our data was during confirmation in word search sequences, as shown in excerpt 5.

Excerpt 5. Repetition as confirmation

```
01 A:      ah:: (.) I will going to m. (0.7)
02            I will go::, (0.7) Doitsu?
03            (0.4)
04 D:      m:
05 C:      German?
06 A:→     Ger[man.
07 D:→        [German.
08 B:      why?
```

In these sorts of action sequences, a speaker designs the turn in such a way as to demonstrate that he or she is having trouble accessing some element of the turn-in-progress, in this case, the word "Germany." Evidence available to the recipients includes: delay-markers (such as 'ah::' and 'm.'), turn-internal pauses, same-turn repetition, and stretched vowel sounds ('go::').[3] All of these allow the speaker to delay completion of the turn while still maintaining the floor. When another participant offers a candidate repair, as C does in line 6, the word-searcher often confirms the candidate token by repeating it with falling intonation.[4] This practice works in a similar way to what Schegloff (1996a; 1996b) has called "confirming allusions." Teachers likewise rely on repetition in the third turn of Initiation-Response-Feedback (IRF) sequences to display acceptance of a student answer to a teacher-initiated knowledge check question (Hellermann, 2003).

So what is clear is that repetition by another speaker in next-turn can accomplish a variety of socio-pragmatic functions. However for the remainder of this paper we are going to focus only on situations when repeating elements of the primary-speaker's prior turn demonstrates listenership, such as that shown in excerpt 1. Given that up until now most studies have focused on repetition in L1 talk, the present study aims to look at some ways second language users claim receipt by redoing part of the prior turn.

Background and Data Set

The data have been gathered from three series of oral proficiency tests video-recorded among Japanese learners of English. The participants were 1st- and 2nd-year university students from a variety of faculties who were undertaking weekly classes in oral English proficiency. The data sets are summarized in Table 1.

Table 1. The Data Set

Data code	Group size	Task length	Recordings	Speaking task
4ninST	4	6 minutes	10	Discuss a topic from class, randomly selected just prior to the test.
Fnds	2	5 minutes	8	Discuss an episode of the sitcom "Friends," which was shown in class.
TB	2	4 minutes	8	Discuss a topic from class, randomly selected just prior to the test.

The video files were saved in MPEG format and, after repeated viewings, were transcribed according to the conventions devised by Jefferson (as documented in Schegloff, 2007 and summarized in Appendix 1). Through extensive consultation and careful observation, we gathered and analyzed a collection of 76 instances of interaction in which repetition accomplished acknowledgement/receipt. While this is a sizeable data set, the CA approach does not attempt to establish generalizability on the basis of frequency, but rather aims to undertake a deep descriptive account of the focal interactional practice. Paraphrasing Sacks (1984b:411), ten Have (1999) notes:

1. The ultimate 'results' of CA are a set of formulated 'rules' or 'principles,' which participants are demonstrably oriented to in the natural interactions.
2. The way to arrive at such results is to analyse singular instances, formulate rules, and 'test' these with comparable other instances (p. 135-136).

Our purpose in this paper is to qualitatively explicate how the candidate phenomenon is achieved temporally and sequentially, and to consider what implications this might have for L2 pedagogy. Several key instances in which these learners achieved receipt through repetition are outlined below.

Findings

Receipt through repetition is an interactional practice that is part of a broader discourse pattern in which one speaker (A) is established as the teller and the other (B) as the recipient. The practice is regularly used in conjunction with other forms of minimal receipt token, including minimal "aizuchi-like" responses such as *un*, *mm*, *ohn*, or what Jefferson (1985, p. 4) has termed "passive recipiency." Speaker B uses repetition to signal to A that he/she understands a given element of the prior turn, and is actively following the general flow of the talk. In essence the practice is a way of displaying listenership that is more specific than just "uhuh" or "mm." The sequence of turns we will analyze can be summarized as follows:

Turn 1.	A:	produces an informing or telling, sometimes in a way that invites uptake
Turn 2.	B:→	provides receipt by repeating some element of the turn-in-progress
Turn 3.	A:	(may minimally acknowledge receipt such as by nodding, then) continues turn in progress, or adds a new *turn increment*

Some further examples of the interactional practice can be seen in the following excerpts:

Excerpt 6: 4ninST 9a Spring

```
01 B:      a=
02 C:      =but eh:: New Zealand is (0.6) south ss:
03         area because so it was ah spring.
04 B:→     °spring.°
05 C:      so (1.0) ah a little >cloudy<.
06 B:      oh::.
07 C:      m.
```

Excerpt 7: 4ninST 6a Six minutes

```
01 T:      talk about travel,
02            (0.5)
03 T:      fo::r six [minutes.
04                   [((beep))
05 C:→     °six minutes.°
06 T:      read[y?
07             [((beep))
```

Excerpt 8: 4ninST 5a Hokkaido[5]

```
01 A:      mm.
02 B:      and I buy, (0.4) Hokkaido.
03 A:→     °ah Hokkaido.°
04 B:      s:ome [  (0.6 )] farmer.
05 A:            [((nods))]
```

In each of these excerpts we can see that one speaker is mainly talking and the other is mainly listening. The recipients cast themselves in the role of listener when they repeat some element of the just-prior turn segment, indicated in the transcripts by an arrow. Note that in each of these cases, the repetition is delivered with falling intonation, differentiating it from the repair-initiator we saw in excerpt 3, which was produced with rising intonation, leading the participants to treat it quite differently in the ongoing talk. Although by no means the rule, we found that the receipt-through-repetition turn was often done with lower volume, as depicted in the transcripts enclosed in °degree marks°, further indicating that the recipients see themselves as the non-primary speaker during the repeated portion of the talk.

In excerpt 7, the turn prior to the repetition is incomplete, and the primary speaker goes on to continue speaking afterwards. This is further evidence to show that the interactants understand the repeater to be in the role of recipient, as the primary speaker goes on to complete the turn, often with minimal or no recognition of the recipient repetition.

At other times the receipt-through-repetition comes at a point where the primary speaker's turn has reached a point of possible completion. In these cases, the recipient often accompanies his or her next-turn repetition with one or more other acknowledgement tokens such as *ah*, *oh,* or *yeah*.

As can be seen from the following examples, in these situations the primary speaker sometimes keeps talking by adding a further TCU (excerpt 9), but at other times receipt-through-repetition can also lead to a change in the primary speaker (excerpts 10 and 11). One associated feature of these action sequences is that the receipt-through-repetition is actually a third-turn action, where the first-turn is a question, the second-turn a response and the receipt completes the sequence, as shown in excerpt 9.

Excerpt 9: 4ninST 3a Kobe

```
01 D:      how-how about you?
02 C:      oh. (0.4).hhh er: (0.4) >ghh<
03         my home city is, (.) Kobe.
04 D:→     Kobe, ah:.=
05 A:      =wo:rgh.
06 T?:     [°sugoi°]
             WOW
07 C:      [  e:r. ] (0.6) Kobe, (0.8) >eh?< (0.6)
08         in Kobe. in mountains, [     (0.6)      ]
09                                [((hand gesture))]
10 C:      exist.
11 D:      [ >un. I know.<]
12 C       [(a:t/and/e:to)] (.) my high school
13         i:s (.) e:::r
```

Excerpt 10: 4ninST 4a Trombone

```
01 A:      ah. I play trombone.
02 B:→     [trombone e:h.
03 D:      [oh:::
04 A:      .hehehah
05 D:      in my junior school junior high
06         school (.) I play (.) flute.
```

Excerpt 11: 4ninST 9a Singapore

```
01         (2.0)
```

```
02 A:      >ah< I have been to: Singapore.
03         (0.3)
04 C:→     OH: Singapore.
05         (0.2)
06 C:      how was it?
07         (0.5)
08 A:      eh?
09         (0.2)
10 C:      how was it?
```

As can be seen from the analysis so far, the repeated turn-segment is usually short, and often consists of only one word. Moreover, participants do not repeat just any word. Svennevig (2004) typified the repeated turn segments as "discrete, detailed pieces of exact information" (p. 502). Almost invariably, the tokens have not appeared in the conversation up until that point and appear towards the end of the primary-speaker's turn. The part that gets repeated regularly consists of one or two short elements—particularly proper nouns. In fact, in roughly one third of the instances we analyzed the repeated element was a place name, such as *Osaka* or *Singapore*. In another third of the cases (36%), the repeated element was a word or phrase that was somehow being negotiated in the interaction. This would suggest that receipt-through-repetition plays an important role in displaying the recipient's understanding of a specific term, especially one that is expected to play some role in the development of the topic in progress.

Note that in almost all of the examples so far the recipient repeats the key element smoothly in next-turn without any hesitation or gap. Research into this practice indicates that this is regularly the case with expert speakers (Schegloff, 1996a, 1996b; Svennevig, 2004). However, in the data we looked at the listener often left a slight gap before repeating the key element, as in line 3 of excerpt 11. This delay may be attributed to the fact that the participants were novice speakers of English, and therefore required more time to formulate their responses (see Wong, 2000). However it may equally be a factor of the group dynamics. We noticed this tended to happen more when the students were talking in groups. In multi-party talk the issue of speaker selection becomes more complicated and such inter-turn silences may appear while participants consider which of the three listeners will voice the receipt. Naturally, this is not such an issue in paired conversations.

The excerpts we have examined so far have all included some sort of "new" and therefore "newsworthy" element that gets acknowledged through repetition. Related to this we found that receipt-through-repetition was regularly used when the participants were negotiating meaning, such as when they needed to circumlocute or apply a novel meaning to some word.

Consider the following excerpt, in which the participants have been discussing whether or not they believe life exists on other planets. Speaker A is listing the sort of countries where he thinks UFO sightings are likely to occur.

Excerpt 12: TB8 USA

```
02 A:      a::h I don't sink so. because. (.)
03         the. a::h (1.0) UFO is, (0.3) the
04         location a::h which, UFO is founded.
05 B:      ohn.
06 A:      ah are so limited.
07 B:      o[hn.]
08 A:       [  f]or example, (0.8) USA:?=
09 B:→     =USA.=
10 A:      =o::r, (0.5) England, [or]=
11 B:                            [un]
12 A:      =France,=
13 B:      =un=
14 A:      =or, so- [   on.  ]
15 B:               [((nods))]
16 A:      and so on.
17 B:      ((nods))
```

In this example the receipt-through-repetition comes in line 9. In line 2 speaker A produces a disagreement followed by an account in lines 3 to 6 (UFOs are only found in certain "limited" places). Speaker B aligns to this turn as a recipient, producing minimal receipt tokens (lines 5 and 7) after A's intonationally complete term increments. Speaker A then furthers his account by initiating a list of examples, beginning with *USA* (line 8). Elongation of the final vowel sound and upward intonation indicate that this is the first of a list-in-progress.

It is at this point that B repeats the word *USA*, to enact receipt and indicate to A that he is following. This receipt-through-repetition facilitates the flow of the conversation, signaling to A that B has comprehended the first element of the list and prompting A to continue with his turn. There is no gap between the repeated token and speaker A's next turn, and the subsequent turn increments are produced in rapid succession, as evidenced by the latching and overlap in lines 8-14. As A's list grows longer, the strength of speaker B's receipt tokens decreases from a repetition of the word *USA* (line 9) to minimal receipts in Japanese *un* (lines 11 and 13) and mere nods (lines 15 and 17).

A similar case can be found several turns later in the same conversation in which A reprises his argument that UFOs are only found in 'limited' places. Again, B is acting as the recipient.

Excerpt 13: TB8 Limited

```
28 A:      [and], the mmm, location. which,
29         UFO is found? is:: limited.
30 B:→     limit[ed yah].
31 A:           [ and  ] (0.3) I think. [.hhh]
32 B:                                  [soh.]
33 A:      .hh heh heha
34 B:      mmm.
```

What exactly A means by 'limited' is unclear, at least to us as analysts. In excerpt 12 he seems to be saying that he does not believe aliens have visited Earth because reported UFO sightings generally only happen in 'a narrow subset' of countries (i.e., Western nations). On the other hand, he could also be using 'limited' to mean 'restricted' referring to the notion that UFO sightings are often reported near classified military bases. The point is that A's usage of the word 'limited' is potentially problematic, and there is evidence in lines 28-29 to suggest that he himself sees it this way. This turn is remarkably similar to lines 3-6 in excerpt 10, and so A's "second doing," along with the upwards intonation on 'found' and the elongation of 'is' in line 29 both indicate he is initiating *forward-oriented repair* (Carroll, 2004) in the form of a word search that ends in the same word 'limited.' This seems to indicate that although A himself is not satisfied with the word he has chosen, it is the most appropriate lexical item available to him at that moment.

Be that as it may, B's turn in line 30 is anything but hesitant. Speaker B accompanies his receipt-through-repetition with an acknowledgement token ("yah"), demonstrating to A that he claims to understand what A means and allowing him to continue speaking. And indeed this is what A does, overlapping his ongoing turn in line 31 with B's turn. It would seem that the first half of B's receipt-through-repetition is sufficient for A to recognize the go-ahead signal. The repetition of the key element 'limited' displays that B understands the term and signals to A that no further explanation is needed.

Whether it is part of a list or a potentially problematic term, it seems that the turn prior to a receipt-through-repetition is sometimes designed to invite uptake from the listener. Let's examine another instance where the repeated element of the turn is offered as an example. In this conversation, A has just told B that he is sometimes attracted to older women.

Excerpt 14: Fnds2 Kuroki Hitomi

```
07 A:      a:::r (1.4) for example? (1.2) °a::h°
08         (1.5) I like, (0.8) >Kuroki Hitomi.<
09 B:→     oh. Kuroki Hi[tomi. ah]-on.
10 A:                   [ a : :h ]
11         (0.3)
12 A:      e::r Japanese actress.
13         (0.4)
14 B:      a(h)h-a(h)h [(h)°oke°] hha [ho.
15 A:                  [ e : : r]     [s:he iz.u?
16         (0.9) fo(r)ty?
```

Here the repeated element is the name of a middle-aged actress who is known for her beauty. The way the initial turn in the sequence is produced has a number of features in common with the earlier excerpts, including turn-internal pauses, elongations, and fillers prior to the eventual utterance of a key turn-final element. In this case it may indicate a word search, as did 'limited' in excerpt 12, or it could be designedly reticent, given that the youthful speaker is admitting he admires an older woman. What is important is that A is introducing a new topic to the conversation, one that B should be expected to either (1) recognize, or (2) clarify, such as by asking, "Who's that?" Without this sort of feedback, A would have to proceed with the topic in a different way, perhaps with a try-marker (Sacks & Schegloff,

1979) or by initiating a confirmation check.

As it happens, B does recognize the reference and expresses this in line 9 with receipt-through-repetition in combination with a turn-initial change-of-state token, *oh*. Heritage (1984) found that such markers demonstrate a change of epistemological states. That is, by saying '*oh*' B is claiming that he has achieved a new knowledge state: he did not know that A likes Kuroki Hitomi, but now he does. He follows this with repetition as receipt, specifying the key element of the news and completing his turn with a further acknowledgement token in Japanese ('*ah-on*'), which may also project agreement (See Ikeda 2007 for discussion of Japanese change-of-state tokens).

As was also the case in excerpts 12 and 13, speaker A's next turn comes quickly, overlapping speaker B's receipt in line 10. *Self-selection* at this point is a *turn-competitive bid* by A, and again B aligns as recipient by waiting in line 11 and listening while A produces a turn increment in line 12.[6]

Repeat, Receipt, and Repair

So far we have noted that doing receipt-through-repetition specifies some element of the prior turn that the listener claims to understand, and that the teller has designed to be somehow 'worth understanding' due to its projected import for the topic-in-progress. The next case demonstrates some of the difficulty experienced in determining whether or not an instance of repetition is intended as receipt—an issue for the participants in real time as much as it is for us as analysts (see Schegloff, 1996b).

Here the same participants from excerpt 14 are discussing whether age makes a difference in a relationship, and A, claiming an extended turn as the primary speaker, begins an account of why he thinks a younger man can love an older woman. While our analysis is only concerned with the first part of this account, the upshot of A's full utterance is that "relationships don't necessarily work out even when the couple is the same age."

Excerpt 15: Fnds2 Wo- wimmen

```
08 B:      O:(gh):[h?
09 A:             [eh: for example e::::h.(2.0) a MA:n?
10 B:      o[hn].
11 A:→      [an]::d.o:: a wo- wimmen?
12 B:→     mm. >[woman. woman.<
13 A:           [e:h there is ((nods))
```

```
14          there are two. >eh< and? (.)
15          a:h they are? same. age.
16          (0.5)
17 A:       [(  )]
18 B:       [o:h.]oh.oh okeh.
```

As in several of the earlier excerpts, just prior to the point at which the listener provides the receipt-through-repetition, the primary speaker is conducting forward-oriented repair. Typical of such word-searches, the first turn in the sequence (line 9) displays uncertainty with a pause, fillers, and elongations, as well as rising intonation. Speaker B reacts to A's rising intonation with a minimal response in line 10 and A continues his turn in overlap (line 11), producing a vowel-marked token[7] ('*and.o*') with considerable elongation at an incomplete point in the turn, which projects a yet-to-appear *trouble source*. Speaker A finally produces the token he was searching for by the end of line 11, first in what is audibly on the way to the singular 'a woman' and then in the plural 'women.' Since it completes the second part of a standardized relational pair (Sacks, 1972) and is grammatically type-related to '*a man*' in line 9, A's first token was actually correct, but for some reason he performs a cut-off before the token is complete, rapidly enacting backwards-oriented self-repair by replacing it with 'women,' such that the token sounds something like '*wo-wimmen.*' Note that while A has 'repaired' his utterance, it wasn't really 'broken' in the first place, and further repair becomes relevant as a possible next action.

However in fact what B does next at the start of line 12 is a minimal response token 'mm,' which may have been triggered by the rising intonation at the end of line 11, as it was in lines 9 and 10. Speaker B then follows this with a swift double utterance of the correct form of the trouble source, *woman*. It is difficult to know whether B intended this as correction or as receipt of A's aborted first try, but by examining the third turn (line 13) we can say that A initially treats it as receipt by carrying on with the sentence, just as the speakers did in the earlier excerpts we examined. Only at the end of his turn does A address the possibility that B was correcting him, by giving a nod, which could be interpreted as either an acknowledgement of the other-repair, or perhaps 'receipt-of-receipt.' At any rate, this seems to be one of the kinds of post-overlapped responses that Jefferson (1993, p. 3) refers to as "attention on the way to something else," as it gives only the barest of acknowledgements before continuing on with the remainder of the turn.

Svennevig (2004) notes that such receipts after 'broken starts' may help the primary-speaker to produce the turn-in-progress by displaying the listener's current interpretation of what is being said, which ultimately accomplishes communication. By formulating the repeated turn as a receipt rather than a direct repair initiator, the recipient accomplishes 'embedded correction' (Jefferson, 1987) of the kind that is regularly used by teachers and expert speakers. Svennevig claims that such receipts make relevant the "linguistic asymmetry of the parties, and constitute a practice whereby native speakers display their construal of [an imprecise] utterance" (2004, p. 504). Our study has confirmed that the same sort of practice can occur between 'nonnatives,' suggesting that the difference is more about relative linguistic expertise than about 'nativeness' per se.

Our final excerpt is an extended sequence of talk in which several instances of repetition appear. A close examination will reveal the ways that this interactional practice not only establishes recipiency but also enables the recipient to take a more active role in co-completing the telling. Taken from the same data set as the previous excerpt, B is attempting to express his opinion concerning age difference in relationships. At this point in the talk, B has stated that relationships with a significant age gap will be unsuccessful, and he is giving examples of some famously mismatched Japanese couples to illustrate his position.

Excerpt 16: Fnds2 Break break

```
18 B:      may- (1.4) uh also un: .ss (0.9)
19         because uh (0.3) ah for example
20 A:      yeah
21 B:      eh: .hh (1.4) Ishida Junichi? tsk
22 A:      ah yeah
23 B:      Hasegawa Yue=
24 A:      =yeah
25            (0.7)
26 B:→     um <bre:ak.hh> ((hand gesture "separating"))
27 A:→     >break break<
28            (0.5)
29 B:      eh: (0.4) Ohsumi Kenya Koyanagi Rumiko,
30 A:      ah yeah
```

```
31              (0.4)
32 A: →      break
33 B: →      break
34           ((laughs))
35              (0.7)
36 B:        uh:: (0.5)
37 A:        Kuroda Arthur [Adachi Yumi break]
38 B:                      [Adachi Yumi break]
39           (0.5)
40 B:        so: (0.2)
41 A:        ah[h:
42 B:          [uhm in Japan (0.3) .ss (0.2) uh
43           (1.4) eh (0.8) age.i problem.u
44 A:        ah[::::
45 B:          [is big
46 A:        yes
```

This sequence is similar to excerpt 12 (USA) in that they both involve listener repetition in the co-construction of a three-part list. In conversation, recipient feedback through repetition seems to be an integral element of expressing lists, which are regularly constructed over multiple turns even by expert speakers of English (Jefferson, 1990).

Speaker B puts forward the names of the first couple (lines 21 and 23), and Speaker A provides a minimal receipt token ('yeah') for each. Speaker B then produces the word *break* (line 26) in a somewhat hesitant manner. Beginning with the noticeable gap in line 25, there is ample evidence to suggest that B appears uncertain about whether the word *break* is an acceptable English term to describe the *breaking up* of a couple: the turn-initial filler, the measured and deliberate way in which he produces the term itself, the mid-word vowel elongation, the turn-final breathing-out, and the co-occurring hand gesture all imply that B is designing the turn as potentially problematic. Even so, A provides immediate acceptance of the term by rapidly repeating it twice in line 27.

It is important to note that the English word *break* does exist as a loanword in Japanese and can be used to describe *a break-up*. These participants have undoubtedly brought this L1 lexical knowledge with them to the conversa-

tion. Together they suggest three examples of couples who have broken up, each following a basic proto-grammatical pattern that could be schematized as [Name1 Name2 break], but appears to be understood by the interactants themselves as [Name1 and Name2 broke up].

Although somewhat slower[8] than those in other excerpts we have examined, B completes the sequence by continuing the list in the third turn (line 29) self-selecting to name another couple as the next example. In line 30, A gives a minimal go-ahead response and briefly waits for B to continue in line 31. However, when B does not complete the turn in a timely manner, A self-selects to produce the word *break* according to the same form that B used in lines 21-26, and this time it is B who repeats the key element in line 33. However, repetition in this case would not be an accurate description: given his slower speaking pace it is more likely that B is simply completing the turn he began in line 29 and A was able to project the appropriate turn ending and produce it before B. Therefore we do not consider lines 32 and 33 as a case of receipt-through-repetition, but instead a co-completion (Lerner, 2002).

At this point the participants switch roles: A, who was mainly the listener, becomes the primary speaker, while B, who has been giving the examples, becomes the recipient. In line 37, Speaker A then proposes a third pair of names to add to the list, allowing B to chorally co-complete the turn (Lerner, 2002) as they name the second partner and produce the word *break* once again, but in unison.

This excerpt reveals how even speakers with limited knowledge of the L2 can engage in conversation without letting their linguistic limitations get in the way. They are capable of co-constructing a conversational sequence even when they themselves have questioned whether their word usage is appropriate. Once the speaker has received confirmation that the listener has accepted the term-in-use, he or she can continue his turn, and even co-complete it with the listener.

Discussion and Conclusion

By this point, some readers are no doubt beginning to wonder why we need to go into so much detail about what seems like such an insignificant thing. In response to such a claim, we would point out that anatomists have documented the tiniest aspects of the human body, naming the parts and describing their functions. As language professionals, we believe that we should be equally interested in the focus of our work, which in this case is

unscripted conversation. The CA approach challenges us to look at mundane talk as a series of socio-pragmatic actions, and to develop an empirical description of the practices of which it consists.

This paper has documented one important interactional strategy available to novice speakers of English. A primary-speaker introduces a new element to the conversation during a telling or an informing. Since it is the initial appearance of this element, recipient recognition becomes relevant as a next-turn action. Sometimes the primary-speaker designs the newsworthy element as potentially problematic but important for the ongoing talk, yet the listener repeats the token in a faster, less problematic manner, often combined with some embodied display of agreement or acknowledgement such as a nod. This receipt-through-repetition provides a brief uptake that signals the primary speaker to continue with the turn-in-progress. In this respect the second turn is similar to other receipt tokens in that it does not imply that the recipient is going to take an extended turn.

Yet this form of receipt does more than just that. It displays a recipient claim to specific comprehension of some key element in the prior turn, allowing the primary-speaker to continue, and to go on using the repeated element in subsequent talk, having established *intersubjectivity* in this temporal and sequential context. Therefore, receipt-through-repetition is an important interactional resource L2 learners can use when negotiating meaning.

While these conversations were all taken from peer-matched oral proficiency tests, there is little in the data to indicate that the practice of repeating a prior-turn segment to enact receipt is limited only to test-talk. Indeed, its use in several of the situations we have looked at would seem to indicate that the students are orienting to the negotiation of meaning. The data in excerpt 5 for example, eventually leads to the negotiation of the word "Germany," and we have noted the students' use of embedded correction in *wo-wimmen* (excerpt 15). By acknowledging a specific token, receipt-through-repetition claims understanding of it at that point, and therefore frees the primary-speaker to progress the topic further. It therefore plays an important role in encouraging second language learners to experiment with language and encourage communication.

One of the authors' initial observations as we discussed the data was that these repeats somehow "sounded Japanese" to us, as if the participants were carrying over Japanese recipiency practices into their L2. Certainly the work by Clancy et al. (1996) suggests that the practice is far more frequent in Japanese than in English. However, the same study also measures the number of

receipts-through-repetition among expert speakers of English so it is obviously also a practice that can be used by so-called 'natives.' Therefore the beginning learners we studied are in fact able to make use of a native-like interactional practice. Our original impression may have less to do with the fact that the participants are Japanese, and more to do with the fact that they are novice speakers of English, and are therefore participating in conversation that requires specific receipt tokens more regularly.

Although the general CA aesthetic emphasizes similarities rather than differences, Wong (2000) notes that novice speakers of English do not use same-turn repetitions to accomplish the resumption of some prior thread of talk after a parenthetical sequence in the way that expert speakers do. She suggests that this skill is something that might be beyond their linguistic proficiency. The kind of repetition that we have looked at in this paper may be the flip side to this argument. Although expert speakers can also use repetition to demonstrate recipiency, they do not seem to do so with the same frequency as the novice participants in the data we examined. Again, this might be partly because there are fewer occasions when an expert speaker does not know a word, so they do not have to produce it in a way that receives repetition.

Given that the initial turn in the sequence often contains some sort of 'designedly unsure' or 'response worthy' item, it is perhaps somewhat natural that novice speakers will need to produce more of these sorts of turns, both due to their limited L2 repertoire and the fact that they are designing the turn for an audience who may not understand. Recipients, in turn, will respond to this with repetition in accordance with the practice we have outlined. While the current study has focused on novice-novice data, Svennevig (2004) has suggested that other-repetition is also often used in novice-expert pairs, such as when instructors repeat a word used by a learner in order to encourage him or her to continue talking.

In fact, in a recent issue of the *JALT Journal*, Sato (2007) has suggested that novice learners of English use this receipt practice more when speaking with experts. Although the current study has not aimed to measure frequencies, its fine-grained analysis may help to reveal some of the reasons behind this phenomenon, and we would suggest that this is more likely to occur because the speakers see a perceived need to make specific their turn-by-turn intersubjective understanding.

While too many of these sorts of repetitions might make the talk seem unnatural, it did not seem to hinder the novice speakers in our data set. If anything, it seemed to help keep the conversation flowing. This suggests that next-turn other-repetition, particularly of problematic words, may help

learners to facilitate communication by providing the speaker with explicit feedback on their moment-to-moment understanding. Teachers and textbook writers would do well to make sure that examples of this practice appear in second language learning resources. Research into shadowing (e.g., Murphey, 2001) has found that the practice of repeating and revoicing turn-final items may in fact help recipients to better comprehend the speaker's speech. Although the participants in this study were not specifically taught shadowing in the classes that led up to these oral proficiency tests, it is possible that some of the students independently developed and made use of a variety of this communication technique.

It is hoped the current analysis will help language teachers appreciate that students possess existing interactional competencies which can enable them to participate in classroom activities and regulate primary-speaker utterances by the way they respond to them. Close attention to the sequential unfolding of this L2 talk has outlined one way that novice speakers display momentary turn-by-turn understandings. Receipt tokens, such as repetition of prior turn segments, provide important cues for the primary speaker, and are therefore consequential for the way that the remainder of a turn-in-progress is designed. Language use and language acquisition are both socially accomplished, and so the best place to look for evidence of learning is in micro-social actions, such as receipt. When all is said and done, in order to make judgments about internally achieved notions like 'learning' or 'cognition,' teachers, testers, and analysts have only what the participants themselves have to go on—external, real-time claims to understanding in conversation. The interactional practice of receipt-through-repetition is one resource for making such judgments.

Tim Greer is an associate professor at Kobe University. His doctoral dissertation uses Conversation Analysis and ethnography to examine ways that multiethnic Japanese teenagers accomplish aspects of their identity in bilingual interaction. He is currently researching naturally-occurring interaction between novice and expert speakers of English.

Vivian Bussinguer is in the Doctoral program at Kobe University. Her dissertation focuses on immigrant children's L1 maintenance and L2 acquisition. She is also interested in bilingualism, heritage language, language loss, and minority language education.

Jeff Butterfield completed a Master's degree at Kobe University in 2007. In addition to his interests in Conversation Analysis and language teaching, he is also researching the textbook issue and its effects on Sino-Japanese relations.

Agnes Mischinger completed a Master's degree with the School of Cultural Studies and Human Sciences, Kobe University in 2007, specializing in non-verbal communication. Her thesis looked at acculturation and adjustment among foreigners living in Japan.

References

Carroll, D. (2004). Restarts in novice turn beginnings: Disfluencies or interactional achievements? In R. Gardner & J. Wagner (Eds.), *Second language conversations* (pp. 201-220). London: Continuum.

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*, 355-387.

Firth, A., & Wagner, J. (1997). On discourse, communication and (some) fundamental concepts in SLA research. *The Modern Language Journal, 81*, 285-300.

Gardner, R. (1997). The conversation object mm: A weak and variable acknowledging token. *Research on Language and Social Interaction, 30*(2), 131-156.

Gardner, R. (1998). Between speaking and listening: The vocalisation of understandings. *Applied Linguistics, 19*(2), 204-224.

Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. Amsterdam-Philadelphia: John Benjamins.

Gardner, R., & Wagner, J. (2004). *Second language conversations*. London: Continuum.

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

Greer, T. (2007) Accomplishing identity in interaction: Codeswitching practices among a group of multiethnic Japanese teenagers. Unpublished doctoral thesis, University of Southern Queensland, Australia. Available online at http://adt.caul.edu.au

Hellermann, J. (2003). The interactive work of prosody in the IRF exchange: Teacher repetition in feedback moves. *Language in Society, 32*, 79-104.

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

Ikeda, K. (2007). *The change-of-state token a in Japanese language proficiency interviews.* Paper presented at the 6th Annual JALT Pan-SIG Conference, Sendai, Japan. Available online at http://www.jalt.org/pansig/2007/HTML/Ikeda.htm

Jefferson, G. (1985). Notes on a systematic deployment of the acknowledgement tokens 'Yeah' and 'Mmhm'. *Papers in Linguistics, 17*(2), 197-216.

Jefferson, G. (1987). On exposed and embedded correction in conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organization*. (pp. 86-100). Clevedon: Multilingual Matters.

Jefferson, G. (1990). List construction as a task and resource. In G. Psathas (Ed.), *Interaction competence* (pp. 63-92). New York: Irvington.

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

Kendon, A. (1967). Some functions of gaze directions in social interaction. *Acta psychologica, 26*, 22-63.

Lerner, G. (2002). Turn sharing: The choral co-production of talk-in-interaction. In C. Ford & B. Fox & S. Thompson (Eds.), *The language of turn and sequence* (pp. 225-256.). Oxford: Oxford University Press.

Markee, N., & Kasper, G. (2004). Classroom talks: An introduction. *The Modern Language Journal, 88*(4), 491-500.

Maynard, S. (1989). *Japanese conversation: Self-contextualization through structure and interactional management*. Norwood, NJ: Ablex.

Murphey, T. (2001). Exploring conversational shadowing. *Language Teacher Research 5*(2) 128-155.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/ dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge: Cambridge University Press.

Richards, K., & Seedhouse, P. (Eds.). (2005). *Applying conversation analysis*. London: Palgrave Macmillan.

Sacks, H. (1972). On the analyzability of stories by children. In J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 325-345). New York: Holt, Rinehart and Winston.

Sacks, H. (1984). On doing "being ordinary". In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 513-529). Cambridge: Cambridge University Press.

Sacks, H., & Schegloff, E. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 15-21). New York: Irvington.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50*(4), 696-735.

Sato, M. (2007). Social relationships in conversational interaction: A comparison of learner-learner and learner-NS dyads. *JALT Journal, 29*(2), 183-208.

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

Schegloff, E. (1996a). Confirming allusions: Towards an empirical account of action. *American Journal of Sociology, 104*, 161-216.

Schegloff, E. (1996b). Issues of relevance for discourse analysis: Contingency in action, interaction, and co-participant context. In E. Hovy & D. Scott (Eds.), *Computational and conversational discourse: Burning issues—an interdisciplinary account* (pp. 3-38). Heidelberg: Springer.

Schegloff, E. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.

Schegloff, E., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361-382.

Schegloff, E., Koshik, I., Jacoby, S., & Olsher, D. (2002). Conversation analysis and applied linguistics. *Annual Review of Applied Linguistics, 22*, 3-31.

Seedhouse, P. (2005). Conversation analysis and language learning. *Language Teaching, 38*(4), 165-187.

Stivers, T. (2004). "No, no, no" and other types of multiple sayings in social interaction. *Human Communication Research, 30*(2), 260-293.

Svennevig, J. (2004). Other-repetition as display of hearing, understanding and emotional stance. *Discourse Studies, 6*, 489-516.

ten Have, P. (1999). *Doing conversation analysis: A practical guide*. London: Sage.

Wagner, J., & Firth, A. (2007). Second/foreign language learning as a social accomplishment: Elaborations on a reconceptualized SLA. *The Modern Language Journal, 91*(Focus issue), 800-819.

Wong, J. (2000). Repetition in conversation: A look at "first and second sayings". *Research on Language and Social Interaction, 33*(4), 407-424.

Yngve, V. (1970). On getting a word in edgewise. *Papers from the sixth regional meeting, Chicago Linguistic Society* (pp. 567-577). Chicago, IL: Chicago Linguistic Society.

Appendix A

Transcription conventions

SIMULTANEOUS UTTERANCES

`huh [ oh ] I see`	Left square brackets mark the start of overlapping talk
`[what]`	Right square brackets mark the end of an overlap

CONTIGUOUS UTTERANCES

`=`	Equal signs indicate that: a) Turn continues at the next identical symbol on the next line, or b) Talk is latched; that is, there is no interval between the end of prior turn and the start of next turn

INTERVALS WITHIN AND BETWEEN UTTERANCES

`(0.4)`	Numerals in parentheses mark silence, in tenths of a second
`(.)`	A period in parentheses indicates a micropause (less than 0.1 sec)

CHARACTERISTICS OF SPEECH DELIVERY

`hhh hee hah`	indicate laughter or breathiness
`g(h)et ou(h)t`	used in parentheses when the laughter occurs within a word
`.hh`	indicates audible inhalation
`hh`	indicates audible exhalation
`mine`	Underlining indicates marked stress
`found? is`	A question mark indicates rising intonation
`yes.`	A period indicates falling intonation
`so,`	A comma indicates low-rising intonation, suggesting continuation
`HUH`	Capitals indicate increased loudness
`°thanks°`	Degree signs indicate decreased volume
`>limited<`	Inward-facing indents embed talk which is faster than

	the surrounding speech
`<break>`	Outward-facing indents embed talk that is slower than the surrounding speech
`go:::d`	One or more colons indicate lengthening of the preceding sound. Each additional colon represents a lengthening of one beat
`wo- wimmen`	A single hyphen indicates an abrupt cut-off, with level pitch

COMMENTARY IN THE TRANSCRIPT

`((hand clap))`	Double parentheses indicate transcriber's comments, including description of non-verbal behaviour
`the (park)`	Single parentheses indicate an uncertain transcription

OTHER TRANSCRIPTION SYMBOLS

→	An arrow in the transcript margin draws attention to a particular phenomenon the analyst wishes to discuss

Appendix B

A Glossary of CA Terms Used in this Paper

Intersubjectivity	The process by which interactants establish shared meanings and understandings in conversation.
Receipt tokens	A short utterance that indicates a listener is following some prior element of a primary speaker's talk. Typical English examples might include "uhuh," "hmm" or "yeah."

Repair	Any of a number of interactional practices speakers use to deal with trouble in talk. *Backward-oriented repair* seeks to rectify some mistake or problem that has already appeared in the talk, e.g., "now (1.1) I don't study (0.4) **dent-** about (0.4) dentistry," while *forward-oriented repair*, such as a word search sequence, addresses trouble that has yet to be made explicit in the turn, e.g., "I live i:n, (0.7) um (2.3) eas- um? (0.5) **eastern part** of Sapporo." Repair can be initiated and/or completed either by current-speaker (*self-repair*) and/or next-speaker (*other-repair*). For further details, see Schegloff, Jefferson and Sacks (1977).
Self-selection	A turn-taking practice in which the speaker elects to speak next (or to continue speaking), as opposed to *other-selection*, which often happens through questions or other sequence initiating actions. See Sacks, Schegloff, and Jefferson (1974).
Trouble source	Any part of a turn that the participants orient to as "in need of repair." This may be, for instance, a grammatical error, a mistaken referent (such as calling someone by the wrong name), a less-than-true statement, or indeed anything that the speakers treat as repairable. For example, in the following sequence, the trouble source is "much": A: much snow. in Kitami. [(0.6)] D: [((moves head toward A))] A: much snow. (0.8) D: no less snow.
Turn-competitive bid	When two or more speakers attempt to start speaking at the same time.

Turn increment	Part of a turn-at-talk, especially one that is yet to be completed.

Endnotes

1. The current study does not focus on native speakers of English, but transcripts from other researchers (Clancy et al., 1996; Gardner, 2001; Schegloff, 1996a, 1996b) lead us to believe that English experts also use repetition to accomplish these actions.
2. Italicized terms are explained in further detail in the glossary (Appendix B).
3. In this case the speaker is also able to specify the target of the search in L1. For further discussion of word searches in bilingual interaction, see Greer (2007).
4. Note that in this case the candidate repair is not actually correct, but the fact that the participants treat it as correct at this point in the talk is all that matters in terms of the understanding they are trying to arrive at. In fact, just after this excerpt, the students negotiate the word further, and eventually C offers "Germany" as an alternate.
5. Here the speaker is talking about his desire to one day buy a farm. Therefore his utterance here should be understood as something like, "I will buy some farmland in Hokkaido." The fact that the turn is ungrammatical and includes some inaccurately used vocabulary only serves to point out that listener B is acknowledging only one element of the turn-in-progress, the word 'Hokkaido.'
6. It is interesting to note that A's turn here is a try-marker, which would normally indicate that he believes B does not recognize the referent he has used, despite the fact that B has made a bold claim to such recognition in his prior turn. The fact that this conversation is happening as part of an oral English proficiency test may partly account for A's action here, either as a bid to include the English-speaking tester (who is co-present but not actively participating in the conversation), or as part of a normative practice in EFL classes by which non-English referents and cultural artifacts are elucidated when first used. There is evidence in the transcript, however, to suggest that B finds A's clarification superfluous. After a noticeable gap in line 13, B responds with laughter, a further on-record receipt token ('okay') and the multiple saying of the Japanese

receipt token 'ah.' Stivers (2004) notes that such multiple sayings occur in sequential environments in which the primary speaker has provided too much information. Their occurrence at this point in the talk demonstrates that, for B at least, the receipt-through-repetition should have been a sufficient signal to A that he understood the referent in question.

7. See Carroll (2004) for further discussion on how Japanese learners of English use vowel-marking in forward repair.
8. This difference in pace could be due to the participant himself. This student generally speaks relatively slowly throughout the complete conversation.

Power-Sharing Between NS and NNS Teachers: Linguistically Powerful AETs vs. Culturally Powerful JTEs

Kyoko Miyazato
Hakuoh University

This study investigates team teaching (TT) relationships between AETs (Assistant English Teachers) and JTEs (Japanese Teachers of English) focusing on power-sharing in Japanese high schools. From September 2003 to March 2004, a naturalistic case study was conducted with two TT pairs during bimonthly visits through class observation and individual interviews. Supplementary data were also collected by interviewing students. Research results revealed that the AETs were given full autonomy because of their language power, which caused the JTEs to become peripheral participants. This resulted in the JTEs' dissatisfaction with their TT performance. Furthermore, the JTEs' identity influenced by language power inequality was deeply involved in their peripheral participation, which was supported by the belief in the native speaker fallacy, the idea that NSs are automatically the best teachers of the language (Phillipson, 1992), at the educational, societal, and individual levels.

本研究は、日本の高校における日本人英語教師 (JTE)と英語指導助手 (AET) のティームティーチング(TT)における関係について、教師間の力配分に焦点を置き、調査することを目的とする。2003年9月から2004年３月まで、2組のTTペアを対象としたケーススタディーが実施され、月２回の訪問時に授業参観とTTペアへの個別インタビューが行われた。補足のデータとして、生徒へのインタビューも併せて行われた。その結果、AETは高い英語力ゆえに授業の自治権を完全に与えられている一方、JTEのTTへの参加は消極的となり、結果としてJTEはTTの出来栄えに対して不満を感じていたことが判明した。また、語学力の不均衡によって影響を受けたJTEのアイデンティティーが、JTEの消極的TT参加と密接に関わっており、このことは、教育界、社会、個人レベルに見られるネイティブスピーカー信仰（ネイティブスピーカーであれば自動的によい語学教師であるとする考え）が一因であることが示唆された。

Keywords: NS-NNS, the JET Program, team teaching, power

JALT Journal, Vol. 31, No. 1, May, 2009

Since the inception of the JET (Japan Exchange and Teaching) Program in 1987, team teaching (TT) involving a Japanese teacher of English (JTE) and an assistant English teacher (AET) has been a distinctive feature of public school education in Japan. Yet team teaching has been dogged by controversy mainly because of team teachers' relationships (McConnell, 2000; Mahoney, 2004; Tajino & Walker, 1998). Naturally, giving up autonomy, one of the basic needs that bring about intrinsically motivated teacher behavior (Deci & Ryan, 1985), must be enormously difficult for teachers who are used to their status as the sole authority in the classroom. Moreover, team teachers in Japan differ from each other in multiple ways—in terms of professional status (teacher-in-charge versus assistant), linguistic proficiency (nonnative versus native speaker) and cultural background (cultural native versus cultural nonnative). These differences are likely to involve power issues. The main purpose of this study, therefore, is to explore TT relationships between JTEs and AETs focusing on nonnative speaker (NNS)-native speaker (NS) power/role-sharing in the classroom.

Literature Review

What is the JET Program?

According to a handbook for JET participants called *The JET Programme 2003-2004*, issued by the Council of Local Authorities for International Relations (CLAIR, 2003), this is one of the world's largest international exchange programs. The program description is as follows:

> The JET Programme enables local authorities (prefectures, designated cities and other municipalities) to employ foreign youth for the purpose of foreign language education as well as promoting international exchange at the community level. By teaching foreign languages at schools nationwide and assisting with international exchange activities organized by local authorities, participants engage in international exchange on a variety of levels with local residents. In this way, the Programme is expected to increase cross-cultural understanding as well as contribute to internationalisation efforts in Japan. (p. 2)

The JET Program has been gradually expanding every year since its inception in 1987. The number of NS participants in 1987 was 848 from four countries, but by 2003 it had reached 6,226 from 40 countries (CLAIR,

2003). As of 1999, there were more than 20,000 alumni of this program (McConnell, 2000).

Jobs for participants in the JET Program are divided into three categories: Coordinators of International Relations (CIRs), who are engaged in international activities in prefectural or municipal offices; Sports Exchange Advisors (SEAs), who are placed with local authorities engaged in sports-related activities; and Assistant Language Teachers (ALTs), who team-teach foreign language classes such as English, French, German, Chinese, and Korean in public elementary, junior, and secondary schools (CLAIR, 2003). However, McConnell (2000) has pointed out that more than 90% of all JET Program participants are AETs, and, therefore, that their primary duty is to teach EFL with JTEs in public school settings.

Historical Background of the JET Program

Compared to the previous language programs such as MEF (Mombusho English Fellow) and BETS (British English Teacher Scheme), which were organized solely by the then Ministry of Education in order to improve EFL education in Japan, the JET Program was originally founded for political reasons. McConnell (2000, p. 1) explained that the JET Program, a $500 million "top-down" project, was a "gift" presented at the "Ron-Yasu" (U.S. President Ronald Reagan and Japanese Prime Minister Yasuhiro Nakasone's) summit in 1986 during Japan's economic boom. In the mid-1980s, Japan needed to deal with the economic conflicts it was having with its business partners, especially U.S.–Japan trade friction. Lincicome (1993, p. 127) stated that Japan's *Kokusaika* was "an action against the criticism of Japan's economic self-centeredness and cultural insularity." Thus, internationalization emerged as a political means of enhancing an understanding of Japan and softening economic criticism against it. Seen in this light, the Japanese government established the JET Program as a means of realizing *Kokusaika* and has, from its inception, hired foreign youth—mostly from Japan's main trading partner, the United States—as JET participants with the expectation that they would enjoy working in Japan and then take those positive experiences with them on completing their assignments. The experiences of these individuals are therefore supposed to have a positive effect on Japan-U.S. economic relationships overall (McConnell, 2000).

As an adjunct to Japan's internationalization, emphasizing the necessity of communicative competence in EFL education was regarded as an important task (Wada, 1994). To implement this task, EFL educational policy has shifted from form-focused instruction based mainly on translating English

texts (*yakudoku*) to Communicative Language Teaching, or CLT (Gorsuch, 1999). For example, in the late 1980s, the Ministry of Education established oral communication (OC) courses in secondary education to develop students' listening and speaking skills. The JET Program was established to promote such changes—English NSs were brought in *en masse* to assist JTEs by providing authentic NS models and opportunities for communication (Samimy & Kobayashi, 2004).

Team Teaching Between AETs and JTEs

Brumby and Wada (1990) defined TT under the JET Program as follows:

> Team teaching is a concerted endeavour made jointly by the Japanese teacher of English (JTE) and the assistant English teacher (AET) in an English language classroom in which the students, the JTE and the AET are engaged in communicative activities. (Brumby & Wada 1990, Introduction, no page number)

Brumby and Wada specified various benefits of TT for students: providing authentic interaction with AETs for learning how to communicate in English; offering a model of interaction with an NS through in-class English conversation between JTEs and AETs, and promoting cross-cultural awareness through the differing viewpoints of the two teachers.

Presumably, TT is beneficial for teachers as well, especially JTEs. That is, another purpose of TT in the JET Program was to create on-the-job training opportunities in order to improve JTEs' English communicative abilities by having them share classes with an NS on a regular basis (Gorsuch, 2002; McConnell, 2000; Wada, 1996; Wada & Cominos, 1994). This was ultimately expected to raise JTEs' awareness of English as a communicative medium and promote CLT in the classroom. Gorsuch reported that AETs have encouraged professional and personal growth in JTEs by exposing them to new and different teaching styles and increasing their communicative English ability.

Present Problems Regarding TT Relationships

TT in the JET Program, however, faces tremendous difficulties and conflict. Tajino and Walker (1998) explained that many of the problems are centered on the relationship between the JTE and the AET. McConnell (2000) also pointed out various power imbalances between the two parties—for example, JTEs' deficiency in English conversational ability, Japanese students'

and society's admiration of NSs, AETs' difficulty in understanding classroom culture, and AETs' exclusion from major decision-making in teaching EFL in Japanese schools.

In implementing the JET program, its designers initially assumed that problems arising from responsibility-sharing would be solved by professional status differences—between the JTE as a qualified and experienced teacher-in-charge, versus the AET as an uncertified assistant, typically with little formal training and teaching experience. However, this planned status difference has caused confusion in the classroom. According to *Resource Materials & Teaching Handbook 2000* (CLAIR, 2000), a major TT-related publication for AETs, for example, the importance of AETs' roles as language consultants and cultural informants is emphasized. However, some AETs were originally used by JTEs as so-called "living tape recorders," based on the assumption that the AETs were only assistants (Kumabe, 1996). Consequently, the AETs' role as an assistant has been questioned and criticized. The criticism indicates that AETs should take a more active role if CLT is to be realized in Japanese EFL education.

In recent years, however, JTEs have tended to take a more passive role, acting as "interpreters" (Iwamoto, 1999; Mahoney, 2004; Tajino & Walker, 1998), which has made team-taught classes more AET-centered. Some JTEs defer to AETs, who are after all NSs of the target language, by surrendering initiative and leadership, owing to feelings of inferiority vis-à-vis their English abilities (Murai, 2004; Tajino & Walker, 1998). Students' perceptions of the AET as the main teacher have also encouraged JTEs to take a less conspicuous role as assistants to the AETs (Iwamoto, 1999).

Another problem with respect to TT involves its legitimacy in Japanese EFL education. Japanese students' primary goal is still to pass entrance examinations wherein English grammar and reading are heavily emphasized (Gorsuch, 1999; McConnell, 2000; Samimy & Kobayashi, 2004; Voci-Reed, 1994). Although TT in the JET Program was established to improve Japanese EFL learners' communicative competence, which was regarded as important for Japan's internationalization, Wada (1996, p. 8) mentioned that TT classes have been displaced from the mainstream goals of English education in Japan under the pressure of grammar-emphasized entrance examinations. In other words, under this pressure, most English classes are conducted in Japanese by JTEs in order to transmit such information effectively. As a result, AETs are not allocated to teach other English classes but only OC courses (Gorsuch, 2002: Mahoney, 2004). Reflecting the lack of consistency of EFL education policy and its implementation, the Ministry of Educa-

tion, Culture, Sports, Science and Technology (MEXT) finally announced in March 2003 that a listening test would be included, beginning in 2006, in the English section of the University Center Examination (*Center Nyuushi*), the preliminary entrance examination used mainly for public universities. However, only a few private universities have a listening component in their examinations. Thus, the discrepancy between the government's directive, aimed at promoting internationalization through CLT, and local priorities with respect to entrance examination preparation has contributed to a loss of legitimacy for TT (Samimy & Kobayashi, 2004).

Reported Difficulties of AETs: Lack of Political Power and Local Language/Cultural Skills

Researchers have pointed out that AETs lack political power as short-term assistants in Japanese schools (McConnell, 2000; Mahoney, 2004; Voci-Reed, 1994). For instance, AETs' appointments are limited in terms of age (they must be less than 35 years old) and length of employment (a maximum of five years).[1] In addition, AETs are not allowed to evaluate or give final grades to students because of their official status as assistants. Thus, AETs have little influence on and involvement in decision making concerning the overall direction of English teaching (McConnell, 2000) and this leads to feelings of frustration and disappointment (Voci-Reed, 1994).

Understanding the local culture of Japanese high schools and the Japanese language seem to present additional difficulties. For instance, researchers have reported high levels of anxiety among Japanese EFL learners in NSs' English-only classes due to the learners' lack of exposure to spoken English (Ellis, 1993), as well as such learners' psychological distance from NSs due to cultural and linguistic differences (Miyazato, 2003). Conversely, McConnell (2000) described AETs' frustration toward one particular aspect of Japanese classroom culture—the lack of responsiveness or shyness of Japanese students. In addition, the dominant Japanese teaching style, which, despite the CLT reforms mandated by the Ministry of Education (see above), still focuses on form rather than meaning and therefore interferes with the implementation of CLT, may prove frustrating to AETs (Browne & Evans, 1994; Ellis, 1996). Thus, both AETs' cultural values and teaching methods may be ineffective because of local cultural realities (Holliday, 1994; Liu, 1999).

Reported Difficulties of JTEs: English Language Deficiency and Native Speaker Fallacy[2]

Some JTEs defer to AETs, who are after all NSs of the target language, because they feel they have inferior English abilities (Kamhi-Stein, 1999;

McConnell, 2000; Tajino & Walker, 1998). According to the newspaper report ("*Sensei ga chikara busokuja*," 2005) concerning a survey conducted by MEXT, only 8.3% of JTEs in junior high schools and only 16.3% of JTEs in senior high school have a TOEIC score of 730 (equivalent to TOEFL score 550) or more.[3] It also reported that only 3.9% of JTEs in public junior high schools and 1.1% of JTEs in public senior high schools conducted English classes mostly in English. These results reveal the reality of JTEs' English language deficiency and their corresponding lack of confidence in conducting classes in English.

In addition, Japanese people's sociocultural image of English and its NSs appears to be deeply involved in JTEs' deference to AETs. The authenticity of NSs' English and an elite or "exotic" image of NS teachers have been noted by Japanese EFL learners (Miyazato, 2003; Sugino, 2002). This attitude is reported to be reinforced by learners' parents, who themselves have doubts about JTEs' English skills (Takada, 2000).

Likewise, various researchers have reported that Japanese people still generally support the supremacy of NS English (Butler, 2005; Kubota, 1998; Samimy & Kobayashi, 2004). The Japanese have historically adopted English and its cultures as a symbol of Westernization, and admire Anglo speakers of English due to their prestigious image (Kubota, 1998; Suzuki, 1999; Tsuda, 1997). Kubota has further argued that Japanese perceptions of the NNS of English as inferior to the Anglo speaker of English make the Japanese wish to identify themselves with white Westerners by learning English.

Power/Responsibility Sharing Between JTEs and AETs

AETs' lower status as assistants was intentionally created in order to equalize the power balance between NSs and NNSs in TT settings (Fujikake, 1996). In addition to the tendency for TT to lead to resistance on the part of JTEs because of their loss of full autonomy in the classroom (McConnell, 2000), JTEs were more hesitant to team teach with AETs at the outset of the JET Program due to JTEs' perceived English communicative deficiency (Ogawa, 1998). That is, it was speculated that AETs would be enormously powerful owing to their language superiority, which would surpass any other advantages that JTEs might have. In fact, AETs' employment conditions (an age limit of 35 and a 5-year employment limit) could ensure that AETs remain politically powerless. In fact, most AETs are recent college graduates in their early 20s (CLAIR, 1992) with little or no formal training or experience in teaching EFL or even teaching itself (Tajino & Tajino, 2000); prior living experience in Japan was limited to a 3-year maximum. In other words,

bringing in young untrained native speakers as assistants was considered less threatening to JTEs, and as such was thought to create more balanced power-sharing in the classroom. Wada (1994) actually revealed that AETs with an equal role to JTEs acted as innovators, which could be perceived by JTEs as creating confusion and friction. Thus, deliberately putting AETs in a lower-status position in the classroom may have been necessary for persuading JTEs, who often have an inferiority complex in regard to their English language abilities, to accept TT.

In sum, power imbalances between JTEs and AETs appear to be caused by the different capabilities of the two parties: AETs with language superiority (language power) and JTEs with political power in the local society and a better understanding of the language learning situation and the learners (political/cultural power). Therefore, examining how the differing power structure in TT pairs influences their role-sharing is a key to understanding TT relationships between AETs and JTEs.

Method

Participants

The participants in this study were two teaching pairs (two paired JTEs and AETs) who were involved in TT in the JET Program at different public senior high schools in the North Kanto district. Team 1 worked for a boys' high school in a small city about 100 km north of Tokyo (School 1). The observed class was a required Oral Communication I (OC I) course containing 40 first-year students. JTE 1, a Japanese male in his mid-40s, had studied in the U.S. during his senior year at university. Based on my observations, JTE 1's English oral/aural skills were high. He had passed the highest level of the STEP [*Eiken*] Test, and his TOEFL score from 20 years previously had been over 600. AET 1, a white American male in his mid-30s, had 4 years of English-teaching experience at the high-school level in the U.S. and was in his 2nd year as an AET. His main stated reasons for applying to the JET Program were his interest in different cultures and the relatively high salary.

Team 2 worked for a relatively new co-educational high school (School 2). The class was a required OC I course consisting of 40 first-year students. JTE 2, a Japanese female in her mid-30s, self-evaluated her communicative English level as "not so good." However, based on my 6-month observation, I found her general English abilities to be higher than she thinks and believe she should be considered as above-average. AET 2, a white American female in her mid-20s, had just started her career and life in Japan. She had taught

high school for 1 year before coming to Japan and was also interested in teaching ESL and getting a TESOL degree in the future.

Qualitative Case Study

Since this study tries to capture the complex reality of TT relationships, a qualitative case study approach was adopted. I followed the two TT pairs in the JET Program for 6 months, collecting data mainly via individual interviews and classroom observations. Case studies have been advocated in educational research (e.g., Johnson, 1992; Stake, 1998; van Lier, 2005; Yin, 2003) as a powerful means of "understand[ing] the complexity and dynamic nature of the particular entity, and to discover systematic connections among experiences, behaviors, and relevant features of the context" (Johnson, 1992, p. 84). Through studying particular phenomena, case studies emphasize the importance of *particularizability*, the opposite of generalizability, which is necessary for investigating a single, particular phenomenon on its own terms, in order to avoid simplification of complex social realities (Ramanathan & Atkinson, 1999; van Lier, 2005). At the same time, case studies provide "comparative information to a wide variety of other cases" (Stake, 1998, p. 198) and assist "readers in the construction of knowledge" (van Lier, 2005, p. 95) regarding the educational phenomenon of interest. A primary purpose of the current research is to stimulate investigation of additional cases in order to understand the dynamics of TT relationships, thereby contributing to the development and improvement of TT relationships in general.

Data Collection Procedures

A naturalistic study using interviews and observations was conducted from September 2003 to March 2004. To begin with, classes team-taught by the pairs were observed during twice-monthly visits. The total number of hours of class observation was 15 hours and observation data were written up in field notes. Individual interviews with the JTEs and AETs were carried out separately. The interviews included general questions about school life as well as specific ones regarding the events that had taken place in the observed classes. The interviews with AETs were conducted in English and those with JTEs were done in Japanese and then translated into English by the author. The AETs, who had a lighter schedule than the JTEs, were usually able to devote about 40 minutes to 1 hour per interview, while the JTEs, owing to other obligations, could spare only about 20–30 minutes. The total

interview time for the AETs was about 9 hours, and for the JTEs was about 5½ hours. The interviews were tape-recorded with the written consent of the interviewees and transcribed for data analysis.

Additional information was collected through interviews with 16 students at the research sites. These interviews were conducted to examine learners' perspectives on TT. The interviews were conducted in Japanese and translated into English by the author.

Results

Results from the main data—the researcher's observations and interviews with team teachers and the students—are here summarized in two subsections: AETs' characteristics and JTEs' characteristics. At the end, satisfaction and role-sharing in TT, a relevant influential factor on TT relationships, is also reported.

AETs' Characteristics

The results seem to support the NS-NNS assumption for AETs: AETs were linguistic experts in the target language but cultural novices in the local culture. That is, the AETs' linguistic and sociocultural power as NSs were perceived, but limited exposure and experience made them lacking in the local language/culture skills and they lacked political power owing to their status as assistants.

Linguistically Powerful in the Target Language

The four teachers interviewed all clearly acknowledged the AETs' linguistic superiority. In particular, AETs' "authentic" English pronunciation was noted by the JTEs. JTE 1 said,

> Students probably want to acquire good pronunciation. That's why I take a more passive role in OC courses and let students hear NSs' English as much as possible. Some JTEs are confident in their speaking abilities, but JTEs are not NSs of English…. I never thought that I would teach OC classes alone. If I had to teach them, I would use a tape recorder, which would definitely make classes boring. Things I can do by myself as an NNS in OC classes are quite limited. (Interview, 10/15/03)

AET 1 supported JTE 1's idea:

> Our goal is to speak as much English as possible, and because I am a native speaker, obviously, my English is better and I think maybe it is more clear for students and JTEs . . . . Well, it's my native language, so I am best suited for it. (Interview, 9/17/03)

Likewise, many students emphasized the importance of learning English from NSs. A student in School 1 illustrated the joy of hearing NSs' authentic English:

> We don't have many occasions to encounter foreigners in our daily lives. OC classes are one of the few opportunities for us to be able to hear NSs' real English. When AETs speak English and I understand them, I become so happy. Why? Because the fact that only English is available motivates us to communicate with them. If Japanese is available, we just depend on that too much. (Interview, 3/3/04)

From a teaching perspective, JTE 2 pointed out the AETs' special abilities—quick recognition and correction of learners' mistakes, lenient attitude toward learners' mistakes, exposure to the target language for students—all of which she thinks come from NSs' high level of grammaticality. However, the AETs in this study were not always linguistically accurate. During my class observations, AET 2 wrote "*excercise*" for "*exercise*." Although spelling mistakes can be considered trivial, it is notable that AETs, as NSs of English, also made spelling mistakes. In fact, Kan (2002) mentioned that quite a few JTEs reported complaints concerning AETs' misspelling as well as their inability to write grammatically correct sentences. He speculated that the government demand for increasing the numbers of AETs has lowered their quality.

Socio-Culturally Powerful in the Target Language

Not only NSs' language superiority, but also their high social image attracted Japanese students' attention. Students voiced their admiration toward AETs. One student in School 2 described his special feeling when he received stickers as a prize from AET 2:

> AETs' praise is special for us. If we got the same stickers from a JTE, we would feel weird. The fact that foreigners acknowl-

> edge our English makes us happy, because it really shows that we are good at English. Besides, American stickers are cool! (Interview, 2/9/04)

AETs themselves acknowledged students' admiration and attention due to their exotic nature as NSs. Because of this over-admiration, in fact, AET 2 worried about JTEs' envy:

> I think that the students tend to like the ALTs, you know, because they are someone different and . . . you know, fun and young.[4] My supervisor told me when I first got here, "Oh, you will be the students' favorite teacher" and all of that. I felt bad for the other JTEs. Well, what do they think? How does that make them feel? Would they resent me because of that? I mean, I haven't experienced any of that here luckily, but I think some of the other ALTs might experience that. (Interview, 1/15/04)

She further mentioned that Japanese students have "a friendly and approachable" image of AETs because of their special position as foreign assistant teachers.

For his part, AET 1 acknowledged the advantages of being a white NS of English. He first mentioned the power of English:

> English is not going away. In fact it's just going to spread. I mean, it's an unfortunate reality that Japanese is not going to become a world language . . . . Those things were in place a couple hundred years ago, I think, for English to do that. Japan obviously is the strongest economy in the world. It's right up there anyway. They want to retain that position. In order to do so, it's English. (Interview, 3/3/04)

He then continued:

> Being born a white, male American, it's like hitting the lottery, in a global sense. I could have been born in poverty in India, just those three things: American, male, and Caucasian. That is the easiest path, or at least one of the easiest. Look at my life here in Japan—I am being paid better than I was as a full-time teacher in America. I am being paid 10,000 dollars more and I am only an assistant . . . . So if you want to be a foreigner in Japan, it is probably best to be a western foreigner. (Interview, 3/3/04)

According to Kan (2002), AETs' monthly salary of 300,000 yen is even higher than that of JTEs of the same age, which is about 180,000 yen.[5] Thus, AET 1 referred to their high salaries as evidence of AETs' special treatment. In return for his high salary, AET 1 explained that he had 18 hours of teaching, but few other obligations. In contrast to JTEs, who have a heavy workload, AETs get paid well and have a relatively light workload in spite of their assistant status.

Thus, AETs' powerful socio-cultural image attracted the attention of not only Japanese learners but also administrators, which seemed to result in their special treatment of AETs. Furthermore, it is assumed that in order to satisfy the students' strong desire to communicate with NSs the JTEs took assisting roles in spite of the fact that AETs' official status is only that of an assistant.

Politically Powerless in the Local Culture: AETs as Assistants

AETs' linguistic and socio-cultural power was not reflected in their political treatment in the educational settings. In other words, despite the fact that AETs were treated as special guests, the results of this study revealed that they still remained politically weak in the education system, since they have the status of "foreign assistants."

In the classroom, AETs were regarded mostly as guests by the students, rather than authoritative teachers who could have a strong effect on students' school work and lives. For example, I observed an incident in which a male student in School 2 cut the TT class and was thereafter scolded by JTE 2. JTE 2 explained,

> The boy cut the class and walked around outside, because I assume AET 2 doesn't scold students. AETs are only guests for students, because they never give them grades. Students just regard AETs as someone that speaks "live" English. In contrast, JTEs give them grades and get involved in student discipline in their daily lives, so students see us as some kind of authority. (Interview, 10/27/03)

AET 2 also admitted that students did not behave well without JTEs' presence. AET 1 likewise explained that the lack of AETs' political power results in their leaving student discipline to JTEs.

In fact, the AET's politically weak position was graphically demonstrated during one of my interviews with AET 1. Two JTEs in the school, who were

strangers to me, came in the room right after knocking quickly and said in fluent English, "This room is reserved for other purposes from now. Can you evacuate now (meaning 'leave immediately')?" Their fluent but direct request sounded like an order to me and I left the room quickly with AET 1. When I asked the AET later how he felt about the incident:

> Some people would say that's rude. I don't see any point in playing on cultural difference. I mean, for me to get upset about it . . . I think don't fight it. Unless they're doing something that I consider immoral or dangerous, I have no intention of telling them they shouldn't do things that way . . . I would expect the same in the reverse situation. I need to listen to them. (Interview, 3/18/04)

He continued:

> I am not going to complain, especially in my situation here . . . . Here I am an assistant. That's another thing. That's another reason why I am not frustrated with any of the teaching methods or why I don't complain, because I am an assistant.

Although this incident may have been caused by the JTEs' lack of understanding of English politeness conventions, AET 1 accepted his position as an assistant, or someone of lower status.

Linguistic and Cultural Barriers

In addition to the fact that the position of "assistants" makes AETs powerless (Voci-Reed, 1994), the results of this study reveal that the AETs are already powerless in Japanese schools due to linguistic and cultural barriers, regardless of their political status.

The language barrier caused by not being conversant in Japanese affected the AETs' teaching significantly, causing frequent confusion in class. For example, although the AETs put great effort into making themselves understood in English by using easy words, expressions, and gestures in the classroom, even simple instructions for activities and games were often not understood well by students. Thus, when AET 1 told the class during one of my observations "After you answer, choose the next person," the students didn't understand. On another occasion, AET 1 asked "Who is the little girl?" in order to find a volunteer to read the part of the little girl in a textbook conversation, but his question confused students, who tried to find the name of

the girl, which was not known.

Although AETs in this study believed that Japanese language skills could help close the cultural and linguistic gap between their students and themselves, AET 1 asserted that mastering the Japanese language requires enormous effort because of its complexity:

> I never really intended, to be honest with you, to learn Japanese all that intricately . . . . Because it is hard . . . . It was a silly idea, but I was thinking that learning Japanese would be like learning Spanish . . . . I mean it is kind of shameful for me to admit this but I really don't study Japanese all that hard . . . . I realize that I am leaving and I realize that I can't master it and [all I want to do is] to really just have fun with it. (Interview, 12/17/03)

As a consequence, AET 1 recognized the importance of JTEs' role in transmitting information to students precisely and effectively:

> I really need a JTE, especially in a writing class, where it is really helpful for a detailed explanation. But even in oral communication, it kind of helps, you know. Maybe I am talking too fast or maybe I am using vocabulary that they don't understand . . . and I mean a lot of people say, that it's best to have just a pure English instruction. I kind of disagree because if you are trying to explain something, just the one push in Japanese by the teacher can lead to so much more understanding on their part. (Interview, 10/15/03)

As for the cultural barrier, both AETs in this study encountered different values from the local society concerning classroom and societal cultures. For example, they revealed their inability to understand their students' silence and passive attitude toward learning. AET 2 said,

> It seems like, especially with Japanese students, because they are so shy and I can basically count on them not volunteering even though I ask for a volunteer. I would like to try just in case a student will raise their hand, but I expect to have to call on the students because I know they are not going to volunteer . . . . I mean, in America, you know, any student is going to get shy or embarrassed by standing in front of their peers, but here it seems a little more extreme. (Interview, 10/6/03)

AET 1 also described difficulties in dealing with silent students:

> I don't like to sit there in dead silence and wait for an answer that is not going to come . . . . Yeah, [students should try to answer] either through a gesture or an expression or a sound. But yeah, that student there, he wasn't moving . . . I knew he was awake. (Interview, 10/15/03)

AET 1 described his changing feelings in this regard:

> But I've gotten used to it. I've gotten to a point where, in the past, answers like, "I don't know" or "No" were unacceptable answers. So, I would press and, [sometimes] just to the chagrin of the student who was being focused on. You see him writhing like, "Leave me alone," so usually I'll back off. (Interview, 2/18/04)

He further explained,

> But, sometimes I feel bad, though, I don't know, maybe if you don't let them answer and move on, well, maybe they were going to answer and somehow they feel like they failed. And even now when that happens sometimes, I'll try to come back to that student later with an easier question. I just get a sense of crushing defeat from that student or something. I'd just like to let them say something, let them think about something to have some small success. (Interview, 10/15/03)

Thus, the AETs still had difficulties dealing with Japanese classroom culture although they knew that Japanese group norms and face issues contributed to the students' passive learning attitude. In sum, the AETs were not only politically but also culturally powerless in relation to the local culture.

JTEs' Characteristics

The results also seem to support the NS-NNS assumption for JTEs: JTEs were linguistic novices in the target language but cultural and occupational experts in the local culture. The JTEs in this study revealed lack of confidence in their English abilities and their belief in the native speaker fallacy. However, the JTEs' important roles as linguistic, cultural, and psychological mediators to fill the gaps between Japanese students and AETs were acknowledged by the teachers and students.

Culturally Powerful in the Local Culture: JTEs as Language/Cultural/Psychological Mediators

JTEs' role as language and cultural mediators was pointed out by the team teachers and students in this study. AET 1 and AET 2 readily admitted that JTEs' language support was indispensable for both the students and the AETs themselves to avoid confusion. AET 2 referred, in particular, to the importance of making complicated directions or difficult explanations more clear. AET 1 also emphasized JTEs' translating role for in-depth understanding and precise explanation. The JTEs themselves acknowledged their role as language and cultural mediators. JTE 2 stated that her role was to prepare a comfortable learning environment by being a gap filler between AETs and Japanese students. JTE 1 also described his role as a language mediator who is empathetic to students struggling with foreign language learning:

> For instance, an easy question becomes difficult for Japanese students because of NSs' authentic pronunciation and intonation. They are not familiar with them. Also, they can be careless and miss NSs' utterances when their concentration breaks down. Their English level is still not so high, so even if they don't understand, they don't know how to say that in English. Learning a foreign language is stressful. (Interview, 3/3/04)

The JTEs' mediating roles were appreciated by students as well. One student in School 1 said,

> We learned that looking away from AETs is an impolite behavior from JTE 1, so I avoid doing so even when I don't know how to answer. It's important to know this, so AETs won't misunderstand us. (Interview, 3/3/04)

Another student in School 2 asserted that JTEs provided psychological relief:

> I can speak with AET 2 without worrying because we have JTE 2. She fills cultural gaps between us and gives us psychological relief. You know, JTEs help us when troubles arise. For example, AETs sometimes misunderstand Japanese students' silence. We may be silent because we are extremely nervous or embarrassed. It's hard for AETs to understand that. JTEs understand Japanese students' feelings better. (Interview, 2/9/04)

In fact, the JTEs' better grasp of Japanese students' lives and feelings were observed during my class observations. For example, JTE 1 paid special attention to students' "face issues" in class. When AET 1 casually asked one student how he liked an activity which involved translating English cartoons into Japanese, the student answered, "So-so." Presumably, AET 1 thought that the student was able to translate the cartoons and asked him to explain it in Japanese for the class, but the student couldn't say anything. The class was uncomfortably silent for a moment. Then JTE1 helped the student translate the cartoon without embarrassing him, by inferring from his facial expression that he actually hadn't understood the cartoon.

In addition, understanding Japanese humor seemed to help create a bond with students. JTE 1 showed his sense of humor with his Japanese students. One day he brought a toy ear to surprise the students. He covered his ear with the toy ear and said "What did you say?" in English. The students burst into laughter.

In particular, JTEs' better grasp of students' vocabulary was acknowledged as one of the major advantages of JTEs. JTE 1 said,

> JTEs know what words are taught by now, so we can easily rephrase difficult words when students don't understand. Moreover, it is a trivial thing, but we know that classes after P.E., for example, should be conducted at a slower pace because students are usually tired after exercising. (Interview, 9/17/03)

Thus, the JTEs' knowledge of the students' lives as well as the local language and culture seemed to create their image as psychological mediators and trust from the students.

Linguistically Powerless in the Target Language

In spite of the numerous advantages, the JTEs in this study pointed out AETs' language superiority. JTE 2 revealed her lack of confidence in her English language abilities. Even JTE 1, a fluent English speaker, confessed that he had been scared of TT when he started his teaching career. He said he had a fear of revealing his inability in understanding AETs in front of students and the AETs themselves, but TT became less of a psychological burden after he gained confidence in his communicative English skills. Even now, he makes efforts to maintain his English skills by talking with AETs, as well as reading about 50 English paperbacks and watching about 70 English movies a year.

JTE 1 further revealed the realities of some JTEs' poor English abilities as follows:

> As a matter of fact, there are still many JTEs who don't want to speak to AETs due to their poor communicative abilities. Actually, I watched a JTE's open class the other day, but to be honest, it was miserable. I know it's impolite to say this, but the JTE's English was terrible. (Interview, 3/3/04)

Moreover, JTE 1 referred to the current emphasis on CLT, which has put the pressure on JTEs to avoid the use of Japanese in class. He explained,

> I have a growing sense of crisis about my status as an NNS in CLT-emphasized policy. Nowadays, English has been introduced in elementary schools. If students acquire speaking and listening abilities in the earlier stages, we will need to teach some kind of content in English. To do that, more NSs will be hired and JTEs who cannot do that may lose their jobs. At this moment, there are still many students who cannot understand English without Japanese explanation, so JTEs are necessary. But in the future, people will recognize the necessity of NSs' authentic English and may criticize JTEs' poor English communicative abilities. Maybe JTEs with near-native abilities will survive, but more training should be given to present JTEs. Otherwise, JTEs will have no choice but to specialize in teaching reading and grammar only. (Interview, 3/3/04)

In sum, it is expected that JTEs' English language deficiency in general results in a lack of confidence, which has led to a situation wherein JTEs were not the total power dominators in spite of their designated political power as teachers, not assistants.

JTEs' "Native Speaker Fallacy"

The JTEs' lack of confidence in English language abilities seems to result in their belief in the native speaker fallacy. For example, JTE 2 confessed her recognition of NSs, especially British and Americans, as ideal teachers of English:

> I know it is a prejudice, but if I were to have formal English education myself, to be honest, I would prefer NSs as my teach-

> ers. For example, if I were to learn English myself and pay for a language school such as AEON or NOVA, I would choose American or British teachers . . . because they speak correct English . . . Of course I know Indian English and Singaporean English are both World Englishes. I have stayed with a Singaporean family in the home-stay program in Canada and it was good training for me to try my English. But if I learn English from them, I cannot help questioning their accent and grammaticality. (Interview, 3/22/04)

JTE 2 explained that students' admiration of NSs led her to take an assistant role in TT settings:

> Our students really look forward to classes with AETs. There are few occasions for them to have them. We only have them once a week, so I hesitate to become the main teacher. I don't want to disturb AETs' classes, so I take the assisting role. I also take the role of student disciplinarian so that AETs can concentrate on teaching and students have a good image of them. Hopefully, this helps to create a fun class atmosphere. (Interview, 11/10/03)

Actually, in the classes taught by both Team 1 and Team 2, the AETs were the main teachers, standing in the center of the classroom, while the JTEs including JTE 1, who has high English communicative skills, took the roles of assisting, translating, disciplining students, and engaging in off-stage chores such as writing on the blackboard and distributing handouts, while standing to the side. JTE 1 expressed a similar opinion:

> Although AET 1 is an assistant and I am supposed to stand in the center as the main teacher, I think it's more natural to position him as the main teacher because we are teaching an English conversation course. Our students also prefer it that way. (Interview, 10/15/03)

Later, JTE 1 clarified his view of English and the NSs as follows:

> It is a fact that English is an international language and we Japanese, speakers of a minor language, are learning the international language . . . . Besides, if NSs were not the lead teachers in TT, inviting them from overseas would be meaningless, anyway. (Interview, 3/10/04)

Thus, the JTEs' belief in the native speaker fallacy influenced their decisions in responsibility/role-sharing in TT settings. To be specific, their belief in the native speaker fallacy seems to be shared with the students and possibly the local people in the society, which might have legitimatized the JTEs' peripheral participation in TT and protected their professional pride as English teachers.

Satisfaction and Role-Sharing in TT

The pairs were asked to give an evaluation of their own performance based on a 100-point scale. The impressionistic question did not have the purpose of rating TT performances statistically. Such measures cannot be used as valid statistics for various reasons, including individual and cultural differences in interpretation of evaluation scales.[6] However, investigating specific reasons for their self-evaluations may reveal their honest feelings about their TT relationships.

AET 1 evaluated Team 1's TT performance as earning 90 points out of 100, saying he had subtracted 10 points for lack of preparation or possible future improvement; AET 2 gave a higher score of 95 points to her team's performance. Presumably, these relatively high scores were because the AET had full autonomy and was the main teacher in class. AET 1 justified the reason for taking the leadership in their TT classes:

> JTE 1 is an *ichinensei* (1st year) homeroom teacher and throughout the week he teaches 16 classes. So they see me once a week. They see him once a day . . . . I mean, of their English instruction, I am a very small part of it. (Interview, 1/17/04)

In fact, JTE 1 admitted that he defers to AET 1, because he can enjoy full autonomy in other courses where he teaches alone. He further revealed his true feelings about TT by giving 75 points to Team 1's performance:

> To be honest, I don't feel like spending time and effort to prepare for only one or two TT classes. OC classes are not as important as classes of grammar and reading. University entrance examinations still emphasize grammar and reading, so OC courses are not main subjects in our school. (Interview, 2/18/04)

JTE 2, who gave 60 points to Team 2's performance, regretted her passive involvement in TT:

> Our students seemed to enjoy our classes because they had many activities and games, which was meaningful and good for the 1st-year students in terms of experiencing an NS's class. But I don't know how much they improved their English. I should have at least gotten involved actively in setting goals and objectives of the course and planning teaching procedures.

Thus, in contrast to the AETs, the JTEs gave lower scores to their TT and did not show full satisfaction in their teaching. Presumably, the JTEs' passive involvement in TT may have led to their lower satisfaction in their TT performance.

Discussion

The results of this study show that while JTEs were linguistic novices, they were cultural experts, and while AETs were linguistic experts, they were cultural novices in this EFL setting. To be more specific, the AETs in this study faced different cultural values as cultural minorities in the classroom due to their lack of cultural power. AET 1 showed low motivation to learn Japanese and AET 2 was a total novice in terms of teaching and living in Japan.

Regarding language power, however, even JTE 1, who had high English communicative abilities, deferred to AET 1, which created a power balance: JTE 1 as the cultural expert and AET 1 as a linguistic expert. However, the power structure did not result in equal role-sharing but led to JTE 1's more passive involvement. This is perhaps why JTE 1 expressed dissatisfaction in their TT performance. Thus, AET 1, as a target language expert lacking political, linguistic, and cultural power in the local culture, became the major teacher in the OC class, which may imply that target-language power could be more significant than any other power in terms of power sharing in the TT setting.

In Team 2, both teachers lacked one of two forms of power—either language or cultural power. JTE 2 clearly revealed her lack of confidence in her communicative English abilities. In other words, while JTE 2 was a linguistic novice, AET 2 was a cultural and occupational novice. However, JTE 2 invariably yielded the floor to AET 2, a 1st-year AET and a total novice in the Japanese school, due to her lack of confidence in her English abilities and her belief in the native speaker fallacy, which may have led to her relative dissatisfaction in their TT performance.

The JTEs' largely peripheral participation was thus deeply influenced by language-power inequality, which seems to be supported by the belief in the

native speaker fallacy at the educational, societal, and individual levels. In fact, as Butler (2005) pointed out, it can be said that the belief in the native speaker fallacy is built into the JET Program itself, based on the following two facts: (1) nearly 96% of AETs are from "core English-speaking countries" (CLAIR, 2003); and (2) they are, by and large, recent university graduates with little or no teaching experience (Tajino & Tajino, 2000). Thus, it is clear that the Japanese government regards NSs from major English-speaking countries, who are, in most cases, not even teachers at all, as qualified and suitable to be assistant English teachers based on their perceived inherent superiority as NSs. Thus, with the initiative of the central government, an influential element in the society, the belief in the native speaker fallacy has become deeply embedded in the EFL profession in Japan.

Concluding Remarks

This study revealed that the particular TT relationships observed turned out to be deeply involved with social complexities of power relations between NSs and NNSs that seem to prominently concern English language proficiency. First of all, it is strongly recommended that JTEs be provided opportunities to establish self-confidence in their language abilities. Surely, the JET Program has contributed to raising JTEs' English communicative abilities through on-the-job training by working with an NS regularly in and outside the classroom. According to *JET Programme: Looking Towards the Future After 15 Years* (CLAIR, 2002), more than 95% of JTEs in the study acknowledged AETs' positive effect on their English abilities. In addition, the central and local governments have recently offered JTEs seminars to improve their teaching and communicative English abilities. However, in order to be confident and comfortable enough to work with NSs, extensive training in the daily use of English is necessary. Studying abroad is also an important option, because it can provide JTEs the additional advantage of cross-cultural experience as well as the experience of being in a linguistic and cultural minority. According to a MEXT action plan entitled "*Eigo-ga tsukaeru nihonjin no ikusei notameno koudou keikaku*" ["Regarding the Establishment of an Action Plan to Cultivate Japanese With English Abilities"] (MEXT, 2003), the National Center for Teacher Development, a government body under MEXT, provided overseas training opportunities for 15 JTEs for 12 months each and 85 JTEs for 6 months each in 2003. In total, then, only 100 JTEs a year have been given the opportunity for intensive overseas training, an enormously small number compared to the nearly 6,000 AETs who have been hired annually through the JET Program. It can also be

speculated that the lack of extensive cross-cultural or overseas experience of JTEs might contribute to cross-cultural misunderstandings between JTEs and AETs.

Another important issue is to lessen or at least call into question the power of the native speaker fallacy in the minds of both AETs and JTEs. First, JTEs as well as AETs should recognize the significance of NNS teachers in EFL settings, who play the crucial role of filling cultural and linguistic gaps between students and NSs based on close familiarity with the learners and experience with teaching and learning skills. Although not investigated in this study, learning/teaching about World Englishes may also lessen students' and teachers' over-admiration of NSs. Through awareness of varieties of English in the world, students may be able to establish appropriate goals and learner identity. In addition, we can consider creating an environment based on Kubota's (1999) critical multiculturalism in Japanese educational settings with various types of EFL teachers. For example, through hiring NNS AETs from different parts of the world and exposing Japanese learners to a variety of Englishes and cultures, learners would be enabled to explore cultural differences without uncritically linking the target language to some exotic culture and get a better sense of and appreciation for World Englishes.

In conclusion, we should carefully consider power issues of English language teaching and learning. Through the insightful divergences such consideration may yield, it is hoped that the Japanese people's belief in the native speaker fallacy will be subject to change.

Acknowledgements

I would like to thank the participants of this study for their cooperation and Dr. Dwight Atkinson for his valuable comments. I also appreciate the useful suggestions made by the anonymous reviewers.

Kyoko Miyazato, Ed.D, is an associate professor of EFL and TESOL at Hakuoh University in Oyama City, Tochigi. Her research interests include team teaching, NS-NNS issues, and intercultural communication.

References

Browne, C., & Evans, B. (1994). The ALT as cultural informant: A catalyst for developing students' communicative competence. In M. Wada & A. Cominos (Eds.), *Studies in team teaching* (pp. 17-28). Tokyo: Kenkyusha.

Brumby, S., & Wada, M. (1990). *Team teaching.* London: Longman.

Butler, Y. G. (2005). *Nihon-no shoggakou eigo o kangaeru—Asia no shiten karano kensho to teigen* [English language education in Japanese elementary schools: Analyses and suggestions based on East Asian perspectives]. Tokyo: Sanseido.

CLAIR (The Council of Local Authorities for International Relations). (1992, May). CLAIR question and answer time. Session conducted at the *JET program(me) Renewers' Conference,* Kobe.

CLAIR. (2000). *Resource materials & teaching handbook 2000*. Tokyo: CLAIR.

CLAIR. (2002). *JET Programme: Looking forward to the future after 15 years.* Tokyo: CLAIR.

CLAIR. (2003). *The JET Programme 2003-2004*. Tokyo: CLAIR.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal, 50,* 213-228.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly, 27*, 91-113.

Fujikake, S. (1996). JET Program to chiiki-shakai [JET Program and its community]. *Modern English Teaching, 9*, 25-27.

Gorsuch, G. J. (1999). *Exploring the relationship between educational policy and instruction in Japanese high school EFL classrooms.* Unpublished doctoral dissertation. Temple University.

Gorsuch, G. J. (2002). Assistant foreign language teachers in Japanese high schools: Focus on the hosting of Japanese teachers. *JALT Journal, 24*(1), 5-32.

Holliday, A. (1994). *Appropriate methodology and social context.* Cambridge: Cambridge University Press.

Iwamoto, K. (1999). Atarashii 'communication' no jugyo niokeru kyodo-jyugyo [Team teaching in new 'communication' courses]. *The English Teachers' Magazine, 3*, 34-36.

Johnson, D. M. (1992). *Approaches to research in second language learning.* White Plains, NY: Longman.

Kamhi-Stein, L. D. (1999). Preparing non-native professionals in TESOL: Implications for teacher education programs. In G. Braine (Ed.), *Non-native educators in English language teaching* (pp. 145-158). Mahwah, NJ: Lawrence Erlbaum Associates.

Kan, M. (2002). ALT-ga fuerunowa iikeredo—Koeni dashite yondewa ikenai ALT mondai [Increasing the number of ALTs is good, but…—ALTs' problems to be discussed secretly]. *The English Teachers' Magazine*, *11, 16-17.*

Kitayama, S., Markus, H., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self. *Journal of Personality and Social Psychology, 72,* 1245-1267.

Kubota, R. (1998). Ideologies of English in Japan. *World Englishes, 17*(3), 295-306.

Kubota, R. (1999). Japanese culture constructed by discourses: Implications for applied linguistics research and ELT. *TESOL Quarterly, 33*(1), 9-36.

Kumabe, N. (1996). ALT donyu ga motarashita mono [What ALTs have brought about]. *Modern English Teaching, 9,* 13.

Lincicome, M. (1993). Nationalism, internationalization, and the dilemma of educational reform in Japan. *Comparative Education Review, 37*(2), 123-151.

Liu, D. (1999). Training non-native TESOL students: Challenges for TESOL teacher education in the West. In G. Braine (Ed.), *Non-native educators in English language teaching* (pp. 197-210). Mahwah, NJ: Lawrence Erlbaum Associates.

Mahoney, S. (2004). Role controversy among team teachers in the JET Programme. *JALT Journal, 26*(2), 223-244.

McConnell, D. (2000). *Importing diversity: Inside Japan's JET Program*. Berkeley: University of California Press.

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2003). *"Eigo-ga tsukaeru nihon-jin" no ikusei notameno koudou keikaku.* [Regarding the establishment of an action plan to cultivate "Japanese with English Abilities"]. Tokyo: MEXT.

Miyazato, K. (2003). Japanese EFL learners' anxiety in native speaker teachers' classes. *JALT Conference '02 Proceedings,* 1-8. Tokyo: JALT.

Murai, G. (2004). Using checklists: A way of reducing surprises in team-taught classes. *The Language Teacher, 28*(11), 35-40.

Ogawa, H. (1998). Kenshu toshiteno JET Program [JET Program as teacher training]. *The English Teachers' Magazine, 1*, 20-22.

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Ramanathan, V., & Atkinson, D. (1999). Ethnographic approaches and methods in L2 writing research: A critical guide and review. *Applied Linguistics, 20*(1), 44-70.

Samimy, K. K., & Kobayashi, C. (2004). Toward the development of intercultural communicative competence: Theoretical and pedagogical implications for Japanese English teachers. *JALT Journal, 26*(2), 245-261.

Sensei ga chikara busoku ja… Kouritsu-chu "Eigo-de jyugyo" wa 4% dake—Monkasho mokuhyo ni tooku [Deficiency of English language abilities of English teachers—Far from MEXT's goal, only 4% of public junior high school teachers conducted English-only classes]. (2005, July 18). *Yomiuri Shimbun*, p. 31.

Stake, R. E. (1998). Case studies. In Denzin, N. K., & Lincoln, Y. S. (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks, CA: Sage.

Sugino. T. (2002) Asia English to gakusei-no ishiki [Asian Englishes and its roles in English education]. *Studies in Humanities and Social Sciences of the National Defense Academy, 85,* 11-31.

Suzuki, T. (1999). *Nihonjin wa naze eigo ga dekinaika* [Why are Japanese poor at English?]. Tokyo: Iwanami Shoten.

Tajino, A., & Tajino, Y. (2000). Native and non-native: What can they offer? Lessons from team-teaching in Japan. *ELT Journal, 54*, 3-11.

Tajino, A., & Walker, L. (1998). Perspectives on team teaching by students and teachers: Exploring foundations for team learning. *Language, Culture, and Curriculum, 11*(1), 113-131.

Takada, T. (2000). The social status of L1 Japanese EFL teachers. *TESOL Matters, 10* (3), 23.

Tsuda, Y. (1997). The hegemony of English vs. the ecology of language: Building equality in international communication. *World Englishes 2000* (pp. 21-31). Honolulu: University of Hawaii Press.

van Lier, L. (2005). Case study. In E. Hinkel (Ed.), *Handbook of research in second languge teaching and learning* (pp. 195-208). Mahwah, NJ: Lawrence Erlbaum Associates.

Voci-Reed, E. (1994). Stress factors in the team teaching relationship. In M. Wada & A. Cominos (Eds.), *Studies in team teaching* (pp. 61-71). Tokyo: Kenkyusha.

Wada, M. (1994). Team teaching and the revised course of study. In M. Wada & A. Cominos (Eds.), *Studies in team teaching* (pp. 7-16). Tokyo: Kenkyusha.

Wada, M. (1996). AET donyu to nihon-no eigo kyoiku [Introduction of AETs and English Education in Japan]. *Modern English Teaching, 9*, 6-8.

Wada, M., & Cominos, A. (1994). Editors' introduction: Language policy and the JET Program. In M. Wada & A. Cominos (Eds.), *Studies in team teaching* (pp. 1-6). Tokyo: Kenkyusha.

Yin, R. K. (2003). *Applications of case study research* (2nd ed.). Thousand Oaks, CA: Sage.

Endnotes

1. The maximum length of employment for AETs used to be 3 years, but has been extended to 5 years under the condition that they are well qualified in regard to their contribution, experience, and motivation.
2. Phillipson (1992) criticized the uncritical belief in the NSs' superiority and called this belief the *native speaker fallacy*—the idea that NSs of English are per se the best teachers of the language.
3. A TOEIC score 730, or TOEFL score 550, is supposed to be the minimum English proficiency needed for international students to enter a university in English-speaking countries. According to a MEXT action plan "*Eigo-ga tsukaeru nihonjin no ikusei notameno koudou keikaku*" ["Regarding the Establishment of an Action Plan to Cultivate Japanese with English Abilities"] (MEXT, 2003), the government expects JTEs in secondary schools to have the score or above.
4. Assistant language teachers (ALTs) include not only AETs but also assistant teachers of other foreign languages such as French and Chinese. However, in common parlance, "ALT" is used interchangeably with "AET."
5. The difference in annual salary between young JTEs and AETs should be marginal when JTEs' bonuses are included.
6. For example, Japanese tend to emphasize self-criticism rather than self-enhancement to fit into the group norm (Kitayama, et al., 1997), and, therefore, they may give lower points in evaluating their TT satisfaction than Americans. It is also expected that JTEs in this study feel reluctant to give a high score on their own performance out of modesty or politeness in front of me, a Japanese university EFL teacher.

Principles and Practices of L1/L2 Use in the Japanese University EFL Classroom

Keith Ford
Waseda University

This article considers the issue of classroom language policy in the context of an EFL environment. Specifically, it focuses on the context of university settings in Japan. Using a semi-structured format, I interviewed 10 university teachers about their principles and practices concerning both teachers' and students' L1 use. Policies ranged from those requiring strict L2 only classrooms to those allowing students to use the L1 whenever they needed. The policies of most teachers were not constrained by any institutional requirements or particularly influenced by critical pedagogy or any language learning theory. Rather, they tended to be determined by pragmatism, individual beliefs, and personality.

この論文では、EFL環境というコンテクストにおける教室内での言語使用の方針について考察する。とりわけ、日本の大学環境というコンテクストに焦点を置く。半構造的インタビューを用いて、教員と学生の第一言語の使用について10人の大学教員に彼らの理念と実践について尋ねた。彼らの言語使用の方針は、第二言語のみの教室を厳密に求めるものから、学生の第一言語の使用を必要である時には認めるものまで、多岐に渡っていた。ほとんどの教員の方針は大学側の要求により強制されたものでも、とりわけ批判的教育学や言語学習理論により影響されたものでもなかった。むしろ、実用的理由、個人的信念、性格により決定される傾向にあった。

Keywords: L1 use, teacher beliefs, classroom practice

The last few decades have seen a growing interest in critical pedagogy in both ESL and EFL contexts. Founded on critical social and educational theory (e.g., Freire, 1972a; Habermas, 1972), critical pedagogy is an approach to language education which places sociopolitical considerations high on the classroom agenda. It incorporates a set of principles

and practices that aim "not only to change the nature of schooling, but also the wider society" (Pennycook, 1990, p. 24). However, while taking a radical critical stance implies emancipatory and transformative action (Freire, 1972a, 1972b; Giroux, 1988), a more practical approach—particularly in an EFL context like Japan—is reflected in calling curriculum and classroom practices into question. In following this tendency, and with particular reference to my own teaching context of tertiary education in Japan, this article addresses the commonly accepted practice of teachers imposing an "English-only" directive on the EFL classroom.

Throughout 10 years of university teaching experience in Japan, I have tended to favor a strict English-only classroom policy, in terms of both teacher and student language use. However, this is something I am beginning to question, from both critical and practical perspectives. Various critical proponents, in presenting their views on *linguistic imperialism* (Phillipson, 1992), the notion of *voice* (Pennycook, 1994) and *participatory* pedagogy and education (Auerbach, 2000), have encouraged a positive reevaluation of L1 use. This has led me to consider to what extent in the EFL context of Japanese universities teachers are making a conscious decision to make use of the L1.

There is also a more pragmatic consideration in making a degree of L1 use a logical alternative to English only: the increasingly low-level English comprehension skills of many 1st-year Japanese university students. In response to rapidly declining numbers of university-age Japanese following decades of falling birth rates, many universities may find themselves adopting a more open policy to student intake by lowering entrance standards. This is reflected in the English ability of incoming students, which at times can create something of a classroom dilemma as many have real difficulty in understanding the most basic native-speaker speech.

In addition to critical and practical considerations, there is a body of research suggesting that L1 use can assist L2 learning at various proficiency levels, such as providing a sense of security to learners and reducing affective barriers (see Auerbach, 1993). Moreover, in terms of theoretical support, Vygotskian sociocultural theory (see Lantolf, 2000), with its concepts of scaffolding, semiotic mediation, and the Zone of Proximal Development, provides an analytical framework supporting student L1 classroom use.

However, despite the support of influential theory and research findings favoring L1 use, much of the anecdotal EFL-teaching literature suggests that English only, as a guiding principle of communicative methodologies, remains a dominant practice among native-speaker teachers. I therefore wished to get a clearer picture of L1/L2 classroom language policies of

native-speaker EFL teachers at Japanese universities, as well as ascertain whether their principles and practices reflect a critical or more practical perspective on language pedagogy.

The L1/L2 Only Debate: A Review

Since the publication of Phillipson's (1992) *Linguistic Imperialism* there has been a radical change in views regarding the issue of the use of L1 in the L2 classroom. Phillipson's work has been particularly influential in critiquing various tenets of the dominant ELT paradigm, such as English being best taught monolingually and by native speakers. This "monolingual fallacy," Phillipson argues (pp. 185-193), is rooted in the maintenance of colonial power and in misguided and negative beliefs about bilingualism. In terms of classroom practices, the imposition of an English-only approach can therefore be considered as authoritarian and reflecting a supposition of linguistic and cultural superiority. On the other hand, students' freedom to use their L1 represents a liberation and democratization of the L2 classroom.

In an ESL context, proponents of critical pedagogy have argued that learners' freedom to use their L1 in the second-language classroom is nothing less than an expression of "linguistic human rights" (e.g., Skutnabb-Kangas, 1999). In such a context, English-only policy is seen as being representative of an ideology maintaining social injustice and existing unequal power relations (e.g., Auerbach, 1993; Corson, 1999). Indeed, Auerbach critiques the tradition of emphasizing *communicative competence* as a primary language goal, suggesting that the resulting practice of "survival English in an English-only classroom" (p. 13) restricts ESL students' participation in society and avoids classroom communication about the real concerns affecting their lives.

In an EFL context, Cook (2001) points out how a dominant English-only agenda has led to the use of such pejorative language as *avoid*, *ban*, and *confess* to consistently describe any L1 use. As a result, for many EFL teachers L1 use is rarely openly discussed, Prodromou (2002, p. 6) observing how it has become something of a "skeleton in the cupboard," a guilty secret. Indeed, some teachers may feel that openly resorting to L1 use represents a contradiction to their traditional direct-method training.

There are, however, strong pedagogical arguments for using L1 in the FL classroom. For example, as Weschler (1997) observes, given the actual time needed to develop any real degree of fluency, limited class time could be better spent on using the L1 as a means of teaching L2 communication skills and strategies. The efficiency argument is further supported by Cook (2001), who suggests L1 use by teachers is more appropriate for task clari-

fication and can lead to more effective learning. Furthermore, in terms of authenticity, Clanfield and Foord (2003) support the appropriate use of L1 in a range of activities, from conversation starters to practicing code switching skills, suggesting that this "reflects the natural interplay of L1 and L2 which is inherent in second language acquisition" (cited online).

Despite strong social, political, and indeed practical arguments for L1 use, caution is needed in simply applying principles of ESL critical pedagogy to EFL classroom contexts. For example, the motivation of an ESL learner who needs English to function in society on a day-to-day basis is likely to be very different to the motivation of an EFL student who is obliged to take an English credit regardless of interest or study major, as is the case at most Japanese universities. Also, it is often pointed out (e.g., Polio, 1994) that in SL settings learners have various opportunities outside of the classroom to develop language proficiency, but in typical FL environments the amount of input is highly limited, with classroom time being learners' only opportunity to actually use their L2 for developing speaking fluency. This is a particular concern in the context of Japanese university English education, where most students have already spent 6 years at high school studying English with the purpose of passing written examinations but have little to no ability to actually use the language in any communicative sense.

One way to approach the L1/L2 only debate from a learner-centered perspective is to find out what students themselves want in terms of L1 and L2 use. For example, in surveying 300 EFL students with Greek as L1, Prodromou (2002) found that while many agreed that the teacher should know and use the L1, paradoxically most "feel they should be hearing and using English" (p. 7), including for procedural uses. Despite encouraging teachers to take a sympathetic view of learners' language and culture, Prodromou concludes from his findings that teachers should not "waste any opportunity to provide students with natural, comprehensible input. Procedural language in the classroom is too good an opportunity for natural English to waste on the mother tongue" (p. 7).

Various university instructors in Japan have also offered results from closed-format student questionnaires to validate the use of Japanese (e.g., Burden, 2000; Critchley, 1999; Cullen & Morris, 2001). Critchley surveyed 160 university students and found that 91% expressed "a preference for some degree of bilingual support in English classes" (p. 13), concluding that "the English-only paradigm may not be entirely appropriate for Japanese contexts" (p. 3). However, such comments and findings often serve to reignite the L1/L2 only debate. For example, in responding to Cullen and Morris, while McAulay (2002, p. 20) calls for foreign instructors to be "displaying

bilingual competence in the classroom, on campus and in the community at large," Ryan (2002, p. 20) cautions against use of the L1, emphasizing that we "maximize the opportunities for students to engage meaningfully in the L2."

But whatever stance is taken regarding this issue, encouraging the use of English in class can be done in a sensitive and rational way. It need not be dictated in an authoritarian manner as in the following anecdotal scenario from Weschler (1997, cited online): "Class. We are here to learn English. As of today, you are not to use any Japanese in this room. This is an 'English-only' class." Such enforcement of an English-only policy may even extend to such extremes as fining students for talking in the L1 and demanding they use only monolingual dictionaries. Critical pedagogues would argue that these measures reflect an agenda of establishing unquestioned control over learners, and directly conflict with their freedom of choice and individual learning preferences. While taking such a dictatorial attitude to classroom language use may be governed by institutional requirements, or by genuine pedagogical principle, it may also result from a teacher's lack of proficiency in the students' L1.

This leads us to consider the key issue of the L1 (i.e., Japanese) knowledge/ability of native-speaker English teachers. Clearly, in a multilingual classroom setting the instructor cannot be expected to know the various languages of the students, and any use of a student's L1 by the teacher may be seen as exclusionary. However, in a monolingual setting like Japan the L1 can be used effectively. Indeed, the main forum for publications by practicing teachers in Japan, *The Language Teacher*, has published various articles addressing the issue of teachers needing a working knowledge of Japanese, including: the call for native-speaker teachers of English to make the effort to learn the L1 to empathize with learners (Barker, 2003); the need for preparing practical and instructional L1 phrase lists (Cole, 1998); the development of sufficient L1 knowledge to make contrastive study part of the language classroom (Yamamoto-Wilson, 1997); and the highlighting of communicative ability in Japanese as being increasingly required for jobs (Glick, 2002).

Given these strong arguments for teachers developing L1 ability for classroom use, it is important to consider a rational and principled approach. Without establishing a set of clear guidelines as to how and when L1 is used it may be difficult for a teacher to monitor not only their students' use but also their own. This has resulted in calls for: the deliberate and systematic use of L1 (Cook, 2001); having "the pedagogical framework" in place to support L1 use (Prodromou, 2002, p. 6); finding a balance that suits one's teaching philosophy (Bawcom, 2002); teachers to be "explicit with regard to activity, purpose,

mode and group configuration" (Polio, 1994, p. 153); and the selective use of L1 "based on critical analysis of their own contexts" (Auerbach, 1994, p. 158).

Despite there being a growing tendency toward recognizing not *whether* to use L1 but *when* to use it, the L1/L2 only issue remains hotly debated. For some, freedom to use the L1 is an ideological issue based on assumptions about existing power relations, while for others a strong commitment to L2 use is based on pragmatism and sound SLA theory that underlines the importance of negotiation for meaning and maximizing L2 input. With this debate in mind, I interviewed 10 teachers concerning their policy toward L1/L2 use.

Method

Participants

The participants in this research were 10 native-speaker EFL teachers, all qualified at least to Master's or Diploma level in TESOL/TEFL or Applied Linguistics. Working at various universities throughout the Tokyo region of Japan, they teach a range of skills and content classes, to both English majors and nonmajors. Their average classes are of about 20-30 students, giving them a clear rationale for using a lot of pair and group work, an important consideration given that I was interested in both teachers' and students' language use. Half of the teachers were female and half male, aged between 30 and 60, and all had spent between 8 and 17 years teaching either at Japanese universities and/or high schools. At the time this research was conducted, they were all colleagues of mine. Five were British, two Australian, two American, and one Irish. Seven of the teachers were part time, working at a number of universities. The other three were full time, either with limited-term contracts or tenured positions. To a certain extent the participants represented a *purposive* sample (Cohen, Manion & Morrison, 2000, pp. 102-104) in that through my knowledge of these instructors I felt they probably represented a range of levels of Japanese language ability and of views and practices regarding L1 use.

Research Questions

The particular research questions that I set out to address were:

1. What stance do native-speaker EFL university teachers in Japan take on the issue of L1 use in the L2 classroom, and how do they rationalize it?
2. Do they set specific guidelines for themselves regarding L1/L2 use?
3. Is their policy negotiated in any way or constrained by institutional demands?

4. Do any of them have a critical perspective to support or question their policy?

Data Collection

As I had specific research questions that I wished to address and that I planned to use as the basis of the discussion section of this paper, I opted for a semi-structured rather than open-ended style of interview. This allowed participants to control topic and introduce various related issues regarding L1/L2 use while still addressing a set of prepared questions that would form the main structure of the interview. Mills (2001, p. 285) describes this kind of interview as one that "combines features of a pre-planned determined framework with scope for the interviewees to talk at length, go off at a slight tangent, pursue a theme."

With two colleagues, I piloted a set of interview questions, changing a number of them, and as a result selecting 14 questions (see Appendix) which acted as a structured framework for further interviews. Though I found that to a certain extent interviewees tended to naturally touch upon questions at various stages in the interview, I did find myself getting them to restate or clarify at the planned point in the interview. I made this a deliberate policy during all other interviews, primarily as a way of facilitating the process of analysis, cross-referencing, and comparison of participants' responses.

Interview Procedure

The interviews took place either at my own home or at locations convenient to the interviewee. Prior to the interview we spent an hour or so chatting over lunch, provided or paid for by the researcher as a means of thanking the participants. Knowing all of the interviewees quite well as teaching colleagues was, I felt, an important contributing factor in creating a relaxed atmosphere for a face-to-face interview.

Before beginning each interview, I stated that I was concerned primarily with classes they taught that had a major oral component, rather than pure composition or reading skills classes. I asked permission to record them using a small unobtrusive portable tape recorder, explaining that only I would be listening to the tapes and that they would be erased once the piece of research was completed. I also clarified the research purpose, for while I had indicated to them prior to the interview that the topic was related to principles and practices of teaching, I had not indicated that the specific topic was L1/L2 use. The main reason for this was that I wanted to ensure a degree of spontaneity and natural flow in talking about their views and experiences.

I was very fortunate that all of the participants seemed to be comfortable in talking about the topic and that they were able to do so with little prompting, encouragement to expand, hesitation, or awkwardness. The fact that they were all keen professionals, and enjoyed their work and talking about it, was undoubtedly helpful. Most interviews lasted 30-40 minutes, with the range being from 25-51 minutes.

Data Analysis

As I listened to each recorded interview, I made summary notes, as well as paraphrasing and transcribing particularly significant comments. This amounted to 41 B5 pages, 3-5 pages for each interview. As I planned to use the research questions as a basis for discussion categories, and in order to assist cross-referencing of topics, I then made marginal notes for each interview using the following lettering system: policy and rationale for teachers' language use (PR-T); policy and rationale for students' language use (PR-S). I also coded comments that suggested a critical perspective (CP); negotiation of policy (NP); and institutionalized policy (IP). Following this analysis I was able to cross-reference and bring together common strands of the 10 interviews for the purpose of presenting results and discussion. In the following discussion, I have used a simple anonymous coding method of P1-P10 to distinguish comments made by the 10 participants (following Cox & Assis-Peterson, 1999).

Results and Discussion

Rationale for Teachers' Use of English Only

In terms of teachers' language use, the majority of participants supported an English-only approach. One strong pedagogical argument for this was that if learners know that the teacher will use the L1, then they stop concentrating so much on processing information in L2, knowing they will get an explanation or instruction clarified in Japanese. Also, as one teacher put it, if students understand that their teacher's Japanese is good, they will use it as a "crutch" (P7) and address the teacher in L1. Knowing they can be more easily understood in Japanese can therefore limit their own opportunities for L2 practice and for developing the ability to negotiate meaning in the L2. One teacher observed that "Once you start speaking in Japanese at all then they don't work so hard, they don't listen as hard, they start thinking more in Japanese" (P5). Furthermore, for those teachers who try to encourage student L2-only use, it was considered that any teacher use of Japanese can act like a knock-on effect: "I would imagine if the teacher is using Japanese, it is giving the go ahead for the students to use Japanese" (P6).

Another rationale for taking a strong line in using only English focused on the issue of teachers recognizing their responsibility to provide students with plenty of rich comprehensible input. A teacher who described himself as "a 100 per cent English-only proponent" (P1) emphasized the need for maximizing students' time spent hearing the L2 as they had so little opportunity to do so otherwise in a typical FL setting ("To be exposed to as much English as possible in a limited time"). He stressed the importance of developing an effective teacher talk style, as opposed to opting to use Japanese for explanation or instructions. It was also pointed out that some teachers may consciously or unconsciously be using classes as an opportunity to practice their own Japanese rather than trying, and possibly struggling, to be understood in English.

One teacher, who was part of a coordinated program with an institutionalized English-only policy, underlined the importance of in-house training and orientation of teachers, as well as students. He suggested this is especially important regarding process language, instruction, and task directions, when the teacher should give "simple instructions, then model" (P2). Furthermore, to facilitate students' L2 understanding they can be easily and effectively trained in various listening strategies and in comprehending instructional, formulaic expressions. This was supported by other teachers who highlighted essential techniques and strategies employed to effectively manage their L2-only policy, such as the careful use of teacher talk, modeling, reiteration, conscious recycling of key phrases and vocabulary, effective use of handouts and board, as well as the need for patience as opposed to using the L1 when students do not understand something the first time.

Rationale for Teachers' Use of L1

While the majority stated that they tried to use English as much as possible, various teachers highlighted the occasional conscious use of L1 for humor and effect, with the purpose of creating a positive, friendly classroom atmosphere conducive to learning. For example, this involved using phrases like "Eigo dake please" ("English only please") and "Wakatta" ("Do you understand?"), as one teacher remarked, simply "to get a laugh" (P1). Another teacher observed, "I use Japanese riddled with mistakes and students just laugh" (P9), the point being not only to create a relaxed and friendly atmosphere but also to communicate the idea that students should not be worried about making mistakes when using the L2.

The occasional use of Japanese by those considering themselves essentially English-only users was also rationalized as an expression of empathy ("I'm trying to learn a bit of your language also" (P3), as a convenient and

simple way of clarifying any confusion regarding instructions or tasks, and as a "time-saving device" (P3), for example using it for short directions as in "Issho ni yomimasho" ("Let's read together"), despite these being—as one strict English-only proponent pointed out—formulaic and easily learnt phrases given practice and patience.

There was, however, one teacher among the interviewees who stood out as taking a particularly proactive stance toward L1 use. He pointed out the initial important goal of students feeling as relaxed as possible in the classroom and of avoiding any possible tension or confusion as a result of not understanding native-speaker English. He considered that teacher use of the L1 was one way of showing support for students dealing with required English courses that they would otherwise find a real struggle: "If you need to speak Japanese to a student to help them get through then that is fine" (P9). The use of "get through" here may reflect how students are seeing English classes primarily as a chore and an obligation rather than as any useful opportunity to develop a skill for their future. Indeed, given that most 1st-year university English classes in Japan are not optional, there is credence in the view that we should assist students in "getting through" these obligatory credits with the minimum amount of stress and discomfort.

Rationale for Students' Use of English Only

The question of if and when students should use the L1 in class appeared to be a more divisive issue than teachers' use. A few teachers expressed a near zero tolerance to L1 use among students, one suggesting that "If your target is a certain language then you should try to get the students to use that language as much as possible. I don't see how there can be any other position really" (P6). She also emphasized that students already have a lot of English from 6 years of English instruction at junior and senior high school, albeit usually highly exam-oriented, and that the university teacher's role is to make them active communicators in the L2 and maximize the opportunities to use what they have: "It's there. It's locked away. It just needs the right opportunity, the right conditions to bring it out" (P6).

One teacher emphasized the need to make an L2-only policy for students' language use absolutely clear from the beginning ("First day I announce this is an English-only class," P10) and to stress that English is a language to be used for the process of "real" communication and not simply as a product for passing exams. In trying to establish an L2-only classroom culture, she extends her policy to all the time that she is actually present in the room. Rather than this leading to silence among students when the class is formally over, they be-

come comfortable with their peers in using English for a "real" communicative purpose: "As I leave I hear them saying to each other *Are you going for coffee?* or *Let's get something to eat*." On commenting on teachers who allow students to use Japanese for the process of naturalistic classroom exchanges, she takes the opinion that "It's telling them that English is not for real communication."

To effectively manage a strong English-only policy regarding students' use of L2, teachers employed a range of techniques and strategies of instruction and classroom management. This involved the teaching of appropriate formulaic phrases and communication strategies that students could use, such as language for requesting and giving clarification, and for conversation expansion. It also involved giving an explicit rationale for an L2-only approach, outlining how L2 maintenance contributed to grading, and the orientation of students in terms of attitude and expectations. Class management involved such strategies as constant pairing up of different students, changing topic often, setting task time limits, and cutting short tasks that proved too challenging for maintaining L2 use. Maintenance of a strong L2-only policy among students also required constant monitoring and verbal reminders or "cajoling and coaxing" (P9), as one participant put it.

Clearly, those who adopt a more learner-centered orientation with high levels of student/student interaction will require considerable management skills and monitoring if they wish to impose and maintain English-only communication among students. Other teachers may restrict interaction to a highly teacher-fronted mode of initiation/response/feedback as a way of controlling and getting students to talk only in the L2, even though the quantity and quality of the communication may be very limited with large classes.

Rationale for Students' Use of L1

There were a few teachers who, while adopting a strong English-only policy for their own language use, were more flexible regarding their students. Examples of when they actively encouraged students' use of Japanese were for the purpose of clarifying among themselves a teacher's instructions and for peer assessment tasks. One teacher, though strongly committed to using only English herself in class, stated "I am much more flexible and permissive about students using L1 among themselves" (P4). She also observed that her attitude toward enforcing L2-only use had changed considerably over the years: "I have softened. At the beginning I was much more of the opinion that it should be English. Speak only English in class. When it didn't work I tried to force it, and that made me uncomfortable."

Her views in particular raise the issue of the nature of classroom content and goals as contributing factors to language use policy. If the type of content is simple personalized day-to-day topics (such as family, food, shopping), L2-only use may not be too demanding. However, if the class content is more sophisticated, requiring critical thinking skills and greater language complexity, then enforcing L2-only use will for many students be counterproductive. Indeed, making this shift from students talking about simplistic, and perhaps, trivial topics and experiences to discussing opinions, values, and ideas related to more challenging issues and content may well be a major influence on language use policy. Not surprisingly, this teacher's view regarding students' use of L1 reflected a high degree of flexibility and goal-orientation. Referring to a content course that she taught, she emphasized that "I wouldn't impose English only on them . . . I would prefer them to speak Japanese" (P4). She supported her approach by emphasizing the importance of plenty of rich L2 input, having respect for students' self-regulation, and the influence of Vygotskian sociocultural theory.

Another teacher implicitly referred to principles of critical pedagogy in supporting students' use of the L1. While a few teachers were unashamedly authoritarian in their use of language (e.g., "I say what I want," P6), and others expressed awareness of their position of power (e.g., "I'm not an English imperialist," P3), this teacher appeared to be particularly sensitive when interviewed to using any language that reflected power and control over students or the exercising of teacher authority (e.g., "I nearly used the word 'control' there, didn't I?" P9). His guiding principles were stated as non-imposition, communication, and the freedom to speak: "I am not going to stop anyone from communicating in class." Students' freedom to communicate in the L1 extended to group project work and preparation of presentations ("I don't really mind how they get there"). He described L2-only policy as "another imposition of I am the boss, I am the one with your grade" and commented on such extreme measures as fining students for using the L1 as being restrictive, controlling, and a reflection of how "the power situation is made very clear, who is in charge and who isn't."

While the above teacher's approach appears to have a clear critical perspective in terms of valuing learners' freedoms and rights to use the L1, it is still essentially a top-down policy determined by the teacher. This leads on appropriately to another critical consideration: that of negotiating classroom language policy.

Negotiation of Policy

One possibility for getting students' views on L1 use is at the beginning of a course to give open questionnaires which may then form the basis for discussion and negotiation of policy. Auerbach (2000) has suggested using awareness-raising model dialogues which can present both sides of the argument and encourage students to discuss their own views on the issue. I was interested to see whether the teachers in this study used negotiation not only as a democratic principle reflecting a critical perspective, but also as a strategy for getting student commitment to L2 communication.

Surprisingly, not one of the 10 interviewees had really used a process of negotiation of language use policy or of specific occasions when L1 and L2 use would be appropriate. In fact, in response to the relevant interview question about negotiation (see Appendix), most teachers offered only very limited comments (e.g., "no, not really" or "no, not explicitly") followed by no further reflection on the issue. However, four of the teachers did offer other views. One teacher recognized that while "negotiation shows respect, gives them investment," his policy was "nonnegotiable" (P8), and he suggested that it would not be a particularly useful process. Another expressed a similar reserve about the efficacy of negotiating policy. In his case, while students attending the university's coordinated English program are asked to discuss why a strong English-only policy is important, the policy is very much a "set agenda," and he considered that once an English-only policy for teachers and students had been established and rationalized as a feature of the program, negotiation of one of its underlying tenets would only prove counterproductive: "What if they come up with, 'Well, we don't have to use English.'? What are you going to do?" (P2).

While negotiation of policy was something they had not tried, two teachers seemed to be more open to considering the possibility. One stated after some hesitation and reflection that he liked the idea and that "I should listen to their opinions" (P1). Another observed "I have been a bit undemocratic in that respect" but suggested that students would probably say that they wanted to speak as much Japanese as possible. She concluded, however, on a more conciliatory note, that "Maybe they would surprise me; maybe it's something I'll try" (P6).

Despite not negotiating their L1/L2 policy, the majority of teachers expressed a strong belief in their approach and that for many of them negotiation could well undermine a policy that they are used to, feel is appropriate, and feel comfortable with. Indeed, in response to the interview question about consistency of policy (see Appendix), most said that they had not real-

ly changed it much over the years, and that it was probably more dependent on their own personality and experiences, as reflected by such comments as: "I'm not like that" (P5), "I'm not very good at being hard line" (P4), and "It's not my style" (P8). As one teacher concluded on this issue, "I've arrived at this point of my principles through intuition and pragmatic reasons" (P3).

Institutional Policy

Most teachers at universities in Japan are not generally constrained by institutional requirements in determining their policy regarding classroom language use. This indeed was the case with nearly all the teachers interviewed. Some observed that in all their time in Japan they had never seen or been given any guidelines regarding this issue ("It is pretty much up to the teacher . . . I have never seen anything explicit," P3). There were two exceptions to this that reflect opposing positions. In the case of one university a change in policy was giving explicit support for teachers to actively use Japanese when formerly it had emphasized staying in English. However, this was only done through a relatively loose directive in a teachers' handbook. This may reflect an increasing trend in the future as the lowering of English standards results in some departments advising English teachers that they can or should use Japanese if students are of such a low level that they cannot understand even the most basic speech in English.

Though most teachers appear to be at liberty to determine their own policies, in a few exceptional cases there may be an institutionally determined policy which demands that teachers and students follow a particular approach to the language used in the classroom, regardless of individual personalities or preferences. Indeed, one of the teachers interviewed was part of a program with a strong institutionalized English-only policy. In-house teacher workshops emphasized the importance of establishing a culture among teachers and students through training and sharing ways to assist in managing classes in a way that values and raises expectations of an L2-only policy. Student orientation included rationalizing the approach to students, such as explaining why an English-only class is beneficial and why Japanese is not to be used.

This example of a strong orientation toward L2-only teacher and student language use has been highly successful in producing and maintaining high levels of motivation and achievement at a university where students are all language majors and have therefore made a very conscious and deliberate choice to be studying English. It is an institution where all the necessary facilities and support are provided to assist their study, including a state-of-the-art self-access centre. In addition, while teachers are expected to commit

to the institution's classroom language policy, they are also encouraged to be involved in research and teacher development. This situation, of course, is in sharp contrast to the situation of most of the teachers in this study, who are taking decisions independently and unilaterally about their teaching practices, without constraints, or indeed support, from a coordinated program.

Conclusion

In reviewing the literature for this study, particularly relating to Japan and other EFL contexts, I found that it was mainly limited to reviewing arguments for and against L1 use and to student questionnaires about preferences. There also seemed to be a general view that English only, as a principle of communicative methodologies, has been the dominant approach in FL settings, though this seems to be based on presumptions about pedagogic traditions rather than on findings from rigorous research. I found no evidence of qualitative, interview-based research concerning teachers' actual policies on this issue at Japanese universities, and so there was a clear rationale for undertaking such a study.

The results showed that 9 of the 10 interviewees tended to follow an English-only approach concerning their own language use. There was a greater degree of flexibility in terms of student language use, with only three teachers claiming a near zero tolerance of L1 use among students. Those teachers who used Japanese occasionally in class clarified the purpose as primarily humor, creating a relaxed atmosphere, giving instructions and task directions. However, none appeared to have established a particularly systematic approach with explicit guidelines regarding the occasions of their own or students' L1 use.

One teacher's L1 use was guided by taking a critical approach to language pedagogy, and another was committed to a strong L2-only institutional policy. Most teachers, however, emphasized the need for a flexible and pragmatic approach, especially concerning student L1 use, depending on student level, motivation, and nature of class content. Above all, teachers underlined the need to feel comfortable with the approaches that they take, with their L1/L2 policies tending to be more the result of personal beliefs and experiences, practical considerations, personality and intuition, as opposed to adherence to any ingrained teaching dogma, training, ideology, or principles of language learning theory.

This study has explored the classroom language policies and the principles and practices of L1/L2 use by only 10 EFL practitioners, and it is limited to the context of Japanese universities. It has provided just one method of undertaking qualitative research in this area. Future qualitative research

might employ longitudinal individual case studies and classroom ethnography over the period of an academic year, involving the triangulation of various data sources, such as classroom observation, teacher and student journals, as well as interviews. Specific areas of research might focus on how L1 use can be systematized, how a strong L2-only policy and L2 classroom culture is established and maintained, or how classroom language policy affects and is affected by teachers' personalities and identities.

Acknowledgement

I would like to give thanks to the two anonymous reviewers whose advice helped me revise this article for publication, and also to Ayako Suzuki for providing the Japanese translation of the abstract.

References

Auerbach, E. R. (1993). Reexamining English only in the ESL classroom. *TESOL Quarterly*, *27*, 9-32.

Auerbach, E. R. (1994). The author responds… *TESOL Quarterly*, *28*, 157-161.

Auerbach, E. R. (2000). Creating participatory learning communities: Paradoxes and possibilities. In J. K. Hall & W. G. Eggington (Eds.), *The sociopolitics of English language teaching* (pp. 143-164). Clevedon: Multilingual Matters.

Barker, D. (2003). Why English teachers in Japan need to learn Japanese. *The Language Teacher*, *27*(2), 7, 9-11.

Bawcom, L. (2002). Over-using L1 in the classroom? *Modern English Teacher, 11* (1). Retrieved September 24, 2007, from http:// www.eltforum.com/articles/free/129.pdf

Burden, P. (2000). The use of the students' mother tongue in monolingual English "Conversation" classes at Japanese universities. *The Language Teacher, 24*(6), 5-11.

Clanfield, L. & Foord, D. (2003). Using L1 in the classroom. *Humanising Language Teachin*g, *5*(1). Retrieved September 20, 2007, from http://www.hltmag.co.uk/jan03/mart2.htm.

Cohen, L., Manion, L. & Morrison, K. (2000). *Research methods in education*. London: Routledge/Falmer.

Cole, S. (1998). The use of L1 in communicative English classrooms. *The Language Teacher, 22*(12), 11-13.

Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review, 57*, 402-423.

Corson, D. (1999). English only and social justice. *TESOL Journal*, *8*(3), 18-22.

Cox, M.I. P., & Assis-Peterson, A. A de. (1999). Critical pedagogy in ELT: Images of Brazilian teachers of English. *TESOL Quarterly*, *33*, 433-452.

Critchley, M. P. (1999). Bilingual support in English classes in Japan: A survey of student opinions of L1 use by foreign teachers. *The Language Teacher, 23*(9), 10-13, 3.

Cullen, B. & Morris, J. (2001). A case for using Japanese in college English classes. *On CUE*, 9(3), 4-9.

Freire, P. (1972a). *Pedagogy of the oppressed*. Harmondsworth: Penguin.

Freire, P. (1972b). *Cultural action for freedom*. Harmondsworth: Penguin.

Giroux, H. A. (1988). *Schooling and the struggle for public life: Critical pedagogy in the modern age*. Minneapolis: University of Minnesota Press.

Glick, C. (2002). Considerations for securing an English teaching position at a Japanese university (Part 2). *The Language Teacher, 26*(9), 19-22.

Habermas, J. (1972). *Knowledge and human interests*. London: Heinemann.

Lantolf, J. (Ed.) (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

McAulay, A. (2002). The paradox of enforced monolingualism for incipient bilinguals. *On CUE*, *10*(1), 19-20.

Mills, J. (2001). Self-construction through conversation and narrative in interviews. *Educational Review*, *53*, 285-301.

Pennycook, A. (1990). Towards a critical applied linguistics for the 1990s. *Issues in Applied Linguistics*, *1*, 8-28.

Pennycook. A. (1994). *The cultural politics of English as an international language*. London: Longman.

Phillipson, R. (1992). *Linguistic imperialism*. Oxford: Oxford University Press.

Polio, C. (1994). Comments on Elsa Auerbach's "Reexamining English only in the ESL classroom." *TESOL Quarterly*, *28*, 154-157.

Prodromou, L. (2002). The role of the mother tongue in the classroom. *IATEFL Issues* (April-May), 6-8.

Ryan, S. (2002). Maximising L2 communication: The case for caution in the use of L1 in the classroom. *On CUE*, *10*(1), 20-21.

Skutnabb-Kangas, T. (1999). Linguistic human rights – Are you naïve, or what? *TESOL Journal, 8*(3), 6-12.

Weschler, R. (1997). Uses of Japanese (L1) in the English classroom: Introducing the functional-translation method. *The Internet TESL Journal*, *3*(11). Retrieved September 20, 2007, from http://iteslj.org/Articles/Weschler-UsingL1.html

Yamamoto-Wilson, J. R. (1997). Can a knowledge of Japanese help our EFL teaching? *The Language Teacher, 21*(1), 6-9.

Appendix

Interview Questions

1. To what extent do you feel you try to follow an English-only policy in the classroom?
2. Do you make your language policy explicit to students from the first few classes?
3. Do you ever consult or negotiate with students about classroom language policy?
4. Can you give examples of when you use the L1 yourself in class?
5. How would you describe your Japanese proficiency?
6. To what extent do you demand and monitor L2 use in pair and group work?
7. Do you include staying in English as part of grading, and if so do students know this?
8. What are some of the techniques you use for keeping your students talking in English?
9. How do you feel when students don't use English and how do you respond?
10. Do you feel you are generally successful at carrying out your policy?
11. What are some of the constraints you find yourself under in carrying out your policy?
12. Is there a general policy at any of the universities where you teach regarding how and when English should be used in the classroom? And if so, do you take that policy into account?
13. Do you feel your policy and attitudes have been consistent over the years or have they changed in some way?
14. Do you think your policy is primarily the result of theory, principle, intuition, experience, pragmatism, or something else?

Revision of a Criterion-Referenced Vocabulary Test Using Generalizability Theory

Takaaki Kumazawa
Kanto Gakuin University

Classical test theory (CTT) has been widely used to estimate the reliability of measurements. Generalizability theory (G theory), an extension of CTT, is a powerful statistical procedure, particularly useful for performance testing, because it enables estimating the percentages of persons variance and multiple sources of error variance. This study focuses on a generalizability study (G study) conducted to investigate such variance components for a paper-pencil multiple-choice vocabulary test used as a diagnostic pretest. Further, a decision study (D study) was conducted to compute the generalizability coefficient (G coefficient) for absolute decisions. The results of the G and D studies indicated that 46% of the total variance was due to the items effect; further, the G coefficient for absolute decisions was low.

古典的テスト理論は尺度の信頼性を測定するため広く用いられている。古典的テスト理論の応用である一般化可能性理論（G理論）は特にパフォーマンステストにおいて有効な分析手法であり、受験者と誤差の要因となる分散成分の割合を測定することができる。本研究では診断テストとして用いられた多岐選択式語彙テストの分散成分を測定するため一般化可能性研究（G研究）を行った。さらに、決定研究（D研究）では絶対評価に用いる一般化可能性係数を算出した。G研究とD研究の結果、項目の分散成分が全体の分散の46%を占め、また信頼度指数は高くなかった。

Keywords: G theory, G study, D study, reliability, criterion-referenced test, diagnostic testing

Classical Test Theory

Classical test theory (CTT) is based on the theoretical foundation that an observed test score is conceptually composed of true score variance and error variance.[1] In other words, the test score variance includes the examinees' true abilities for a target construct, which the test is designed to measure, and measurement error, which creates noise in the testing. The underlying concept of the reliability theory states that if the test succeeds in spreading the examinees' test scores relatively along a continuum or exhibits a large degree of variance, the reliability coefficient is likely to be high. Therefore, the test can be said to estimate their true ability with relative accuracy. That is, the observed test scores vary because the examinees behave differently on the target construct being measured, not because of random noise in the test (Strube, 2000).

Based on the theoretical foundation, actual mathematical formulas were developed to estimate reliability coefficients. The core of the reliability formula is derived by dividing the true score variance by the observed score variance.[2] The most widely reported reliability coefficient is the Cronbach alpha internal consistency reliability formula.[3] If the reliability of a measurement is found to be .80, it indicates that 80% of the observed test score variance represents the examinees' true abilities and 20% is the result of random error creating inconsistency in estimating the examinees' true scores. Such error may be caused by examinee carelessness, testwiseness. or other factors that can result in inconsistency (Brown, 1996). Reliability is also indicative of consistency. For example, sometimes we can draw the inference from a reliability estimate that approximately 80% of the time, the examinees' test scores will vary in the same ways even if they repeatedly take the same test.

In CTT, a standard error of measurement (*SEM*) value for the entire test can also be calculated to show a range within which examinees would probably score if they repeatedly took the same test. Based on the reliability coefficient and standard deviation derived from the test scores, the *SEM* is easy to estimate[4] and interpret. For instance, if the *SEM* was found to be 2.00 and a particular examinee's score was 50.00, the *SEM* indicates that the examinee's test scores would fall between 48.00 and 52.00 about 68% of the time, if the same test were taken repeatedly.

In short, reliability breaks down a set of observed test scores into true score and error variances. However, CTT can only deal with error variance as a single entity and therefore cannot deal with multifaceted sources of error variance. This CTT notion is rather simplistic and not maximally useful

because it is impossible to define the sources of error. In an actual testing situation, numerous facets—such as the number of tasks, passages, and raters—may cause measurement errors. Further, examinees may respond to such facets in complex ways. Therefore, when numerous facets are inherent in a testing situation, the sources of measurement error should be investigated cautiously.

Generalizability Theory

This section introduces the background of G theory and discusses its advantages over CTT. G theory, introduced by Cronbach, Rajaratnam, and Gleser (1963), was extended by Cronbach, Gleser, Nanda, and Rajaratnam (1972) and has been discussed in numerous books on psychological measurement (Brennan, 1983, 2001; Fyans, 1983; Shavelson & Webb, 1991; Strube, 2000; Suen, 1990; Thompson, 2003). This theory was developed as an extension of CTT to investigate the sources of variance in the facets of measurement and to generalize the universe score or true score in CTT obtained from one observation to numerous observations (Brown & Hudson, 2002).

One of the powerful features of G theory lies in the first phase of the investigation called a G study. The multifaceted nature of testing can be broken down into each of the relevant facets of variance, enabling the study of the degree to which the facet variances contribute to the total variance of the test scores. The facets to be examined will depend on the testing situation involved. In performance testing, typical facets include examinees' abilities, rater severities, item difficulties, and occasion difficulties. The variance components for each facet in a particular testing situation can be estimated using an analysis of variance (ANOVA) procedure.

Another advantage of G theory over CTT is that it provides a more adequate estimate of reliability for criterion-referenced tests (CRTs). In CTT, the variability of the test scores is often highly related to the reliability of the test. Since the purpose of norm-referenced tests (NRTs) is to spread examinees' test scores out along a continuum, such variance is appropriate for determining the reliability for NRTs. In contrast, with CRTs, the variance may be suppressed due to three main factors: (a) small sample size, (b) homogeneity of students' proficiency levels, and (c) negatively skewed distributions of test scores at the end of a course. In CRTs, the sample size is relatively small because a limited number of students take classroom-level tests. In a language program where placement tests are administered to create homogeneous classes, that homogeneity is likely to suppress the variance in test scores. Ideally, criterion-referenced items have to be developed

based on class content, such that if all the students learn all the content, they should all score 100% on the test. This can create a negatively skewed distribution that is perfectly logical and as a result suppress the variance. In sum, the CRT's purpose remains to estimate students' achievement in a specific domain. Thus, CTT reliability does not fit the purpose of estimating criterion-referenced measurement consistency; therefore, G theory should be applied to estimate dependability for CRTs. CRT dependability is analogous to NRT reliability in CTT (Brennan, 1980).

Another advantage of G theory over CTT is found in the second phase of the investigation, called a decision study (D study). In CTT, the Spearman-Brown prophecy formula[5] can be employed to estimate reliability with different numbers of items. However, this formula cannot deal with multi-faceted sources of error in a measurement. To estimate the dependability of CRTs in different parallel tests, the index has to be determined based on the multiple sources of error estimated in a G study. The result of a D study is extremely useful in deciding how to revise or redesign a CRT. For instance, let us assume that sections and items are the facets in a given testing situation. The D study allows for calculating the degree of dependability for different hypothetical scenarios, that is, based on different hypothetical numbers of sections and/or items. This constitutes the most practical application of G theory.

A G study should be carefully designed and conducted to investigate the variance components for facets in a given test. Depending upon the testing situation and the measurement design adopted, the study can be designed as crossed or nested and balanced or unbalanced. If all the levels of one facet are the same in the levels of another facet, the two facets are considered crossed. For example, if the five different categories (say Content, Organization, Grammar, Mechanics, and Vocabulary) are scored by three raters, the categories facet is said to be crossed with the raters facet. Alternatively, if all the levels of one facet are different within the levels of another facet, the first facet is said to be nested within the second one. For example, if 10 items in each of three subtests are all different (i.e., items 1-10 are in subtest A, items 11-20 in subtest B, and items 21-30 in subtest C), the items are said to be nested within the subtests.

If all levels of all facets have the same number of observations per facet the design is considered balanced. For example, if all three subtests have 10 items each, it is a balanced design. Conversely, if the levels of even one facet have unequal numbers of observations, the design is considered unbalanced. For instance, in a performance test, if three subtests have different numbers of items (say 8, 12, and 18), it is an unbalanced design.

Based on variance components that can be extracted using an ANOVA procedure in a G study, a G coefficient can be estimated. A G coefficient in G theory is analogous to a reliability coefficient in CTT. Therefore, a G coefficient for norm-referenced (i.e., relative) decisions for a G study design of *p X i* can be estimated by dividing the persons variance component by persons variance component plus persons-by-items interaction variance component (divided by the number of items).[6] True score variance in CTT is analogous to the variance component for persons in G theory, while error variance in CTT is analogous to the variance component for the persons-by-items interaction in G theory. Therefore, G theory is an extension of CTT, but G theory has the additional benefit of making possible the estimation of separate variance components for all possible facets in a testing situation. Under identical conditions, the magnitude of a Cronbach alpha reliability coefficient and G coefficient for relative decisions should be nearly equivalent.

However, G theory can also be used to help in making criterion-referenced (i.e., absolute) decisions based on the extent to which students have mastered a certain domain. In this case, the equation is slightly different from the equation for relative decisions: here, the persons variance component is divided by the persons variance component and items variance component (divided by the number of items) plus persons-by-items interaction variance component (divided by the number of items).[7] The difference between the equations for relative and absolute decisions lies in how error variance is defined. For relative decisions, in the present case, the error variance is defined as the persons-by-items variance component (divided by the number of items). However, in the equation for absolute decisions, error variance includes both the persons-by-items interaction component (divided by the number of items) and the items variance component (divided by the number of items). With NRTs, administrators aim to estimate an examinee's true ability relative to a norm using the test; therefore, the focus is on persons and the interaction of persons with items, and items variance itself is excluded from the equation. However, in CRTs, teachers aim to estimate students' mastery over the item content or domain; therefore, the items variance is included in the equation.

A D study is used to answer a "what-if" question in that it is used to estimate the expected G coefficients if the numbers of items or raters are set at various levels. In other words, a D study generalizes the expected G coefficients under different hypothetical scenarios based on the extracted variance components in the G study. The D study can be conducted by changing the number of items for either relative or absolute decisions. In CTT, after

estimating a reliability coefficient, the Spearman-Brown prophecy formula can be employed to estimate the expected reliability coefficient by increasing and decreasing the number of items in the equation. Although a D study is analogous to the Spearman-Brown prophecy formula, the former can only estimate reliability for changes in one facet (usually items). In contrast, a D study can estimate the expected G coefficients along one, two, or more facets (e.g., items, raters, subtests, and occasions) by setting different numbers of facets at the same time (Suen, 1990).

In the field of educational measurement, numerous articles have been published that apply G theory, particularly for performance testing (Brennan, 2000; Brennan, Gao, & Colton, 1995; Cronbach, Linn, Brennan, & Haertel, 1997). With regard to language testing, only a few books refer to G theory (Bachman, 1990, 1997, 2004; Brown & Hudson, 2002). Brown (1982) first applied G theory to ESP testing. Brown (1993) and Kunnan (1992) investigated CRTs' dependability and employed criterion-referenced item analyses. Lynch and McNamara (1998) applied G theory and the multi-faceted Rasch model to develop ESP speaking tests. They contrasted the two analytical techniques. Employing a large data set from TOEFL, Brown and Ross (1996) and Brown (1999) investigated variance components for the test takers' number, items, subsections, and nationalities. They discovered that the interaction effect caused the most error variance.

In Japanese contexts, few studies have applied G theory (e.g., Yamanishi, 2004). Apart from Griffee's study (1995), which demonstrates the design and evaluation of CRTs using criterion-referenced item analyses, no other study has analyzed teacher-made, criterion-referenced language tests.

Research Questions

In this study, a vocabulary test was developed for a particular class and criterion-referenced item analyses were conducted. What makes the study different is that the test's dependability was estimated by conducting a G study followed by a D study to investigate the optimal number of items and sections needed to achieve a certain magnitude of the G coefficient. In the process, the following two research questions were raised:

1. To what extent is the vocabulary test dependable in terms of the G coefficients for absolute decisions?
2. How many items and subsections are optimal to achieve a certain magnitude of the G coefficient for absolute decisions?

Method

Participants

One hundred thirty-one 1st-year university students enrolled in a required general English course majoring in literature, law, or economics at a high-ranking private university in the Kanto area participated in this study. Four reading and listening classes taught by two instructors were selected. Their goals included improving students' listening comprehension so that they could understand English instructions when taught by native or nonnative teachers in the institution and improving their reading skills and speed. An additional goal included vocabulary development. The teachers set the following goal for vocabulary development: to get approximately 70% of the multiple-choice items correct. The test was designed to gauge the extent to which students learned the receptive meaning of the target words that appeared in the assigned textbook. At the beginning of the first semester, all students were placed into homogeneous groups according to their level.

Materials

To estimate students' mastery of the vocabulary items in the assigned textbook—developed by the English program for a particular course—a vocabulary achievement test was designed and developed. Five chapters were randomly selected from 10 and the items were prepared. Five target words were also selected at random from each chapter in the process of preparing the items; that is, 5 items, from a total of 25, were nested within each section. A sentence identical to one given in the textbook was provided with an underlined target word. All the items were multiple-choice questions, and the choices were written in English. The students were required to choose the answer closest in meaning to the target word. A sample item is as follows:

1. The idea of the need for a common language across the world has become <u>prominent</u> in the twentieth century.

a. Important
b. Common
c. Nonsense
d. Problematic

In the above example, the distractors are *common*, *nonsense*, and *problematic*, which were selected from high-frequency or academic word lists. The test mainly estimates students' receptive knowledge and their ability to gauge meanings from a given context.

Procedure

First, all the textbook passages were scanned and digitalized, following which WordClassifier (Denies, 2004) was employed to classify all the words in the passages in the order of frequency. For each chapter, the target words were selected based on the results of the frequency count. Preceding the test development, test specifications were prepared to clarify the test's purpose and to set a sample test item. During the first week of class, the teachers clearly explained the syllabus, including its goals, objectives, and grading system. While explaining the grading system, they announced that two tests would be administered, at the beginning and end of the course. While the pretest encouraged the students to perform well, it did not affect the students' grades; however, the posttest score accounted for 15% of their final grades. After the procedure was explained, the test was administered in the second week of the second semester in 2005; this was a diagnostic test to gauge the students' knowledge of the target vocabulary items before instruction. The test scores were to be used for the pedagogical purpose of allowing teachers to focus on helping those students with low scores. For vocabulary instruction, the teachers presented a list of vocabulary words for every chapter and provided the Japanese translations and synonyms. An alternative test form was planned to be administered as an achievement posttest for the final assessment.

Analysis

All the items were dichotomously scored, with any missing data treated as an incorrect item. ITEMAN (Assessment Systems Corporation, 1996) was used for the descriptive statistics, distracter analysis, and norm-referenced item analyses such as item facility (IF), item discrimination (ID), and reliability. All the responses were entered into Excel spreadsheet format for conducting criterion-referenced item analyses such as the *B*-index, agreement statistic (A-statistic), and item phi (φ-index). The *B*-index indicates the degree to which a criterion-referenced item differentiates mastery from nonmastery students. The A-statistic indicates the degree to which students answering the item correctly are identical to those who passed the test (Brown & Hudson, 2002). The φ-index essentially refers to the correlation "between examinee item and test performance outcome, their mastery of the item to their mastery of the test" (Brown & Hudson, 2002, p. 126). These statistics are a family of cut-point indices. Based on the cut-point of the test, which was set at 70%, the students scoring higher or lower than 18 were identified as belonging to the mastery or nonmastery groups. XCalibre (As-

sessment Systems Corporation, 1995) is a software program based on the three-parameter logistic model belonging to item response theory. It was used to estimate the KR-21 reliability of the vocabulary test. This software's command file follows the same format as that followed by ITEMAN. Subsequently, GENOVA (Crick & Brennan, 1983) was used to conduct the generalizability and decision (G and D) studies. GENOVA enables users to conduct balanced design G and D studies for random and fixed effects. Here, the G study was a *p* X (*i: s*) balanced design. I treated sections in the textbook as a facet for investigating the extent to which sections variance contributed to the total variance. This design was adopted because five items were nested in each section. After extracting the variance components for all the effects, a D study was conducted to investigate the dependability of the test. Then, the results were processed in Excel spreadsheets.

Results

Table 1 provides the descriptive statistics. The mean of the vocabulary test was 12.37 out of 25; this is desirable because it reveals that the examinees have not yet mastered all the vocabulary words. However, it would have been more desirable if the mean had been lower with a positively skewed distribution. This would reveal that most of the examinees had little knowledge of the target words. Based on the Cronbach alpha and the KR-21, the reliability coefficients for the vocabulary test were found to be .64, indicating that the CRT spread out the examinees' abilities fairly well. Or put another way, the test consistently measured 64% of the examinees' abilities, with the remaining 36% occurring due to error. The *SEM* derived from the Cronbach alpha reliability coefficient was 2.29, indicating that approximately 68% of the time, the examinees' scores would remain in a band that was 2.29 points above or below their observed scores. However, the coefficients and *SEM* are mainly used for interpreting the NRTs' results.

Table 2 summarizes the item analyses. Despite the fact that the IF and ID statistics are norm-referenced item statistics, they provide insightful information for criterion-referenced items. For diagnostic tests, IF values should be low enough to enable students to participate in class and then perform well on achievement tests. For instance, the IFs for items 8 and 18 are extremely high at .80 and .81, indicating that most of the students had already learned the target words before instruction. The mean proportion for the correct items was .49, which is desirable for norm-referenced purposes; however, it would have been more desirable for diagnostic purposes if the value had been slightly lower. Apart from items 8, 18, and 21, which had high

Table 1. Descriptive Statistics for the Total Score

k	*N*	*M*	Variance	*SD*	Skew	Kurtosis	Min	Max	Alpha	*SEM*
25	131	12.37	14.39	3.79	-0.27	-0.45	4	21	0.64	2.29

Notes. Skew = skewness; Min = minimum; Max = maximum; Alpha = Cronbach alpha; *SEM* = standard error of measurement

or low IF values, the remaining items had a large degree of variation. The items with IF values above .50 tend to be negatively skewed; those with IF values below .50 tend to be positively skewed. Most of the items have negative kurtosis values, indicating a flat distribution. Except for items 6, 10, and 20, all ID values were above .20 with a mean ID of .32. In other words, the items discriminated among the examinees' abilities. Ten out of 131 students scoring above the set cut-point were identified as mastery students.

The values of ID and the *B*-index are quite different. In particular, although item 20 was a potential candidate for revision from a norm-referenced perspective, it was a suitable item from a criterion-referenced perspective. Notice that the values of *B*- and φ-indices were nearly equivalent. Items 1, 2, 6, and 25, which have low *B*-index and φ-index values appeared to be problematic. Notice also that the values of the *B*-index and A-statistic are quite different. The A-statistic indicates agreement between answering correctly or incorrectly and passing or failing the test, while the *B*-index indicates the items' capacity to differentiate between students who passed and failed the test. Although item 20 is inappropriate from a norm-referenced perspective, it is suitable from a criterion-referenced perspective because most students who passed the test got this item correct.

Table 3 shows that the items effect and interaction effect accounted for 46% and 52% of the variance, respectively, accounting together for 98% of the total variance. Therefore, the total variance was mainly due to items and interaction effects. The universe score or persons effect included only 2% of the total variance.

A D study was conducted by using the variance components extracted in the G study. In Table 4, the dependability of the vocabulary test with the five sections per five items (k = 25) was found to be .30, which is very low. If the test were to be revised to contain six sections per five items (k = 30), the dependability would be .34, a slight increase. Similarly, if the test were increased to six sections with ten items (k = 60), the dependability would increase by .21. This reveals that a lower number of items and sections results in unsatisfactory dependability.

Table 2. Criterion-Referenced Item Analyses

	Item	IF	Variance	Skew	Kurtosis	ID (Rpbi)	*B*-index	A-statistic	φ-index
1	Prominent	0.35	0.23	0.63	-1.63	0.28	-0.06	0.62	-0.03
2	Guarantee	0.66	0.22	-0.70	-1.53	0.51	0.15	0.38	0.08
3	Emergence	0.39	0.24	0.46	-1.82	0.27	0.34	0.64	0.18
4	Mutual	0.51	0.25	-0.05	-2.03	0.24	0.42	0.55	0.22
5	Diversity	0.68	0.22	-0.78	-1.42	0.46	0.35	0.40	0.20
6	Civilization	0.47	0.25	0.14	-2.01	0.19	0.04	0.53	0.02
7	Ethnicity	0.50	0.25	0.02	-2.03	0.34	0.22	0.53	0.12
8	Clash	0.80	0.16	-1.53	0.34	0.43	0.21	0.27	0.14
9	Scarce	0.55	0.25	-0.20	-1.99	0.36	0.27	0.50	0.14
10	Tremble	0.39	0.24	0.46	-1.82	0.18	0.34	0.64	0.18
11	Equivalent	0.27	0.20	1.02	-0.97	0.32	0.68	0.79	0.40
12	Clinging	0.67	0.22	-0.74	-1.48	0.37	0.25	0.39	0.14
13	Dwelling	0.41	0.24	0.36	-1.90	0.32	0.53	0.65	0.28
14	Excavation	0.38	0.24	0.49	-1.78	0.23	0.45	0.66	0.25
15	Glimpse	0.34	0.23	0.67	-1.58	0.20	0.49	0.70	0.28
16	Restrict	0.65	0.23	-0.63	-1.63	0.58	0.38	0.43	0.21
17	Intimate	0.60	0.24	-0.39	-1.87	0.33	0.22	0.45	0.12
18	Domestic	0.81	0.16	-1.59	0.54	0.47	0.21	0.27	0.14
19	Bury	0.37	0.24	0.53	-1.75	0.35	0.35	0.66	0.19
20	Gullible	0.50	0.25	-0.02	-2.03	0.45	0.21	0.53	0.11
21	Intimidate	0.15	0.12	2.04	2.19	-0.01	0.17	0.82	0.13
22	Distinct	0.29	0.21	0.94	-1.14	0.27	0.23	0.71	0.13
23	Substitute	0.56	0.25	-0.23	-1.98	0.34	0.15	0.47	0.08
24	Sophistication	0.59	0.24	-0.36	-1.90	0.32	0.23	0.46	0.12
25	Ignorance	0.48	0.25	0.08	-2.03	0.22	0.13	0.53	0.07
	M	0.49	0.23	0.02	-1.41	0.32	0.28	0.54	0.16

Note. Skew = skewness; Rpbi = point-biserial correlation

Table 3. Variance Components for the G Study

Source	Variance components	Standard error	Percentage
p	0.003534	0.001016	2%
s	0.000000*	0.006504	0%
i X s	0.091668	0.027881	46%
p X s	0.001102	0.001486	1%
p X i:s	0.104974	0.002910	52%
Total	0.201277		100%

*After Brennan, (1983, pp. 47–48), the negative variance component found for this facet was rounded to zero.

Table 4. Dependability for D Study

	Items									
Sections	1	2	3	4	5	6	7	8	9	10
3	0.05	0.10	0.14	0.17	0.21	0.24	0.27	0.29	0.32	0.34
4	0.07	0.12	0.17	0.22	0.26	0.29	0.33	0.36	0.38	0.41
5	0.08	0.15	0.21	0.26	0.30	0.34	0.38	0.41	0.43	0.46
6	0.10	0.18	0.24	0.30	0.34	0.38	0.42	0.45	0.48	0.51
7	0.11	0.20	0.27	0.33	0.38	0.42	0.46	0.49	0.52	0.54

	Items									
Sections	11	12	13	14	15	16	17	18	19	20
3	0.36	0.38	0.40	0.41	0.43	0.44	0.46	0.47	0.48	0.49
4	0.43	0.45	0.47	0.48	0.50	0.51	0.53	0.54	0.55	0.56
5	0.48	0.50	0.52	0.54	0.55	0.57	0.58	0.60	0.61	0.62
6	0.53	0.55	0.57	0.58	0.60	0.61	0.63	0.64	0.65	0.66
7	0.57	0.59	0.60	0.62	0.64	0.65	0.66	0.67	0.68	0.69

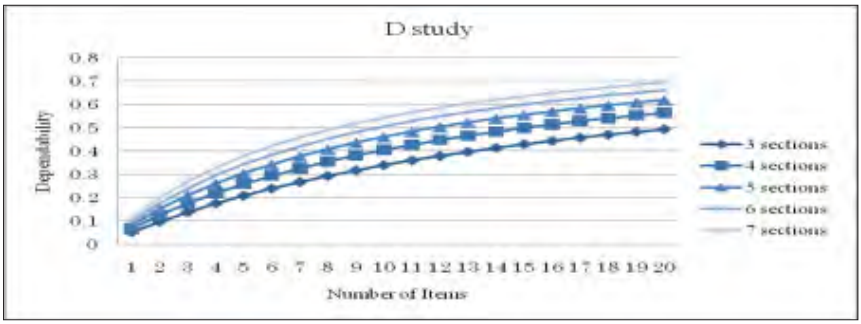

Figure 1. Graphic Representation of the Dependability for the D Study

Discussion

This section discusses the study's research questions, mainly based on the results of the G and D studies.

1. To what extent is the vocabulary test dependable in terms of the G coefficients for absolute decisions?

Two potential reasons for the lack of variability in the persons effect are: (a) sample size and (b) group homogeneity. Nationwide NRTs or placement tests are administered to numerous examinees; however, CRTs are usually administered to relatively small, homogeneous groups of students. The sample size in this study was 131, which is comparatively small from a norm-referenced perspective. Further, examinees with different backgrounds and proficiency levels take NRTs; however, a nearly homogeneous student group, similar in educational backgrounds and proficiency levels, take CRTs. In this study, the test was administered in four classes that two teachers were in charge of. With the exception of one class—identified as a high proficiency group based on a placement test—the proficiency levels of the classes were similar.

Because five items were nested within the corresponding chapters, a G study design of *p X (i:s)* had to be adopted. The results showed that no sections effect was observed. Some students studying only particular chapters of the assigned textbook and not the other chapters might yield sections variance in the posttest score. However, at this time, the students had not studied the textbook. Thus, it was reasonable that no sections variance was observed in this pretest because it did not test how many students had

learned the target words in each chapter. Another possibility was that this multiple-choice vocabulary test was context-independent (Read, 2000). That is, the examinees were able to answer the test items correctly without referring to the context or reading the embedded sentences.

The large variability in the items effect was an interesting result because, thus far, no studies have yielded a similar degree of variability. For NRTs, the persons effect should be large, whereas the items effect should be about one-third less than the persons effect (Brown & Ross, 1996). However, for CRTs, because students are homogeneous in terms of their proficiency level, the persons variance may be low. In addition, since a CRT should be based on items that measure a certain criterion or objective, the large amount of item variance found here may be desirable.

Due to low dependability, the items should be revised. Table 2 shows that the IF values range from .15 to .81. In a diagnostic test, it is desirable that the IF values be generally low, indicating that students have not yet learned the words. The IF values for items 8 and 18 were relatively high compared with the other items, and therefore, they should be excluded from the test. Further, based on the values of the *B*- and φ-indices, the items with low values should be revised. Items 1, 2, 6, 17, and 25 are the candidates for revision. For example, in item 1, students were made to choose the word closest in meaning to *prominent* from the following four choices: (a) important, (b) common, (c) nonsense, and (d) problematic. The correct response is (a). The distractor analysis reveals that the IF values for the four choices were .35, .19, .06, and .40 and the ID values were .28, .42, .03, and .08, respectively. Although ID is a norm-referenced statistic, it can prove useful during the revision of the items. The ID for choice (b) is higher than that for (a), indicating that students with higher scores are more likely to choose (b). The two words, important and common, are synonyms for *prominent*; therefore, both may be correct. However, students with high scores selected (b). Thus, (a) should be replaced with another word so as to function as a distractor and (b) should be the correct choice.

2. How many items and subsections are optimal for achieving a certain magnitude of the G coefficient for absolute decisions?

The Cronbach alpha was moderate, at .64. Although the purpose of this CRT was not to spread students' test score, there was moderate consistency in the test scores. As can be seen in Table 3, the dependability for this CRT was .30. Similar to the classroom tests analyzed in Brown (1993) and Kunnan (1992), this test, too, was not as dependable as expected. First, the G

coefficients for absolute decisions are generally lower than the Cronbach alphas and G coefficients for relative decisions (Brennan, 1980). Second, since these test scores did not affect their final grades, some students may not have taken the diagnostic test seriously; this was a low-stakes test.

Table 2 reveals that the variance component for sections effect was zero. In other words, adding another section to the test would have no effect on its dependability. The results of the D study are presented in Table 4; they reveal that increasing the number of items could contribute to the variability in the students' test scores and produce a higher dependability because a large variability was observed in the items effect. However, the administration time would be longer. In this testing situation, the test should not take over 20 minutes. While developing a test, teachers have to consider dependability and practicality. Finding the "happy medium" (Brown, 1996, p. 34) is the key for revising the test to ensure that it is dependable and practical. The maximum number of items that can be incorporated in the test are 40 because of time constraints in this testing situation. Otherwise, students would not be able to complete the test within the stipulated 20 minutes. If the test contained 40 items, then, based on the D study, the dependability would be .41.

Conclusion

In this study, G theory and criterion-referenced item analyses were applied to revise a CRT. While NRTs are used to spread examinees' test scores out, CRTs are designed to estimate students' mastery of specific objectives or language points. The *B*-index, A-statistic, and φ-index were used for the criterion-referenced item analyses; G theory was also applied to estimate the dependability of the domain score. In addition, the study showed that a G study can capture the multifaceted nature of testing by examining the degree to which the facets (sections and items nested within sections in this study) contribute to the total variance. A D study was applied to determine the optimal numbers of items and sections needed to make the test more dependable and practical in a revised version.

Before developing the test, it is crucial for teachers to thoroughly conceptualize its design in terms of purpose, content, procedure, target domains, number of items, sections, constraints, and analyses. Test specifications (a) are a good way to describe the design, (b) can guide test development, and (c) can serve as the basis for validity arguments to defend the diagnostic or achievement decisions that affect students' lives.

Often, preparing and marking a test is a cumbersome process that causes teachers to lose interest in analyzing their own tests. Teachers who neglect this procedure as part of teaching practice should recognize the importance of learning from the data. Sometimes, the expected result can differ completely from the actual results; therefore, the data analysis should be considered as part of good practices that confirm the extent to which expectations and results match.

The classroom tests must be developed before the actual teaching occurs so as to enable teachers to be aware of what is going to be tested; this will lead to the implementation of successful teaching-to-test instruction with the objective of maximizing students' achievement. Further, diagnostic tests are not often administered as part of teaching practices because the administration of tests takes up class time. However, the results can provide a great source of information, helping to identify misplaced students or mismatches between the students and the class objectives. In this study, 10 students scored higher than the stipulated cut-off based on the diagnostic test administered in the second semester. However, for reasons yet unknown, the students did not perform well in the placement test and were therefore not placed in the correct class levels. It is possible that they effectively learned vocabulary during the first semester or the summer vacation. If a large proportion of students scored above the cut-point, it is possible that the objectives were not set properly. Here, most of the students were nonmastery students; therefore, it was not necessary to change the materials or redesign the objectives.

The result of diagnostic tests can also be used for pedagogical purposes: to identify students' strong and weak points. The teaching should focus on the objectives that were not attained by students to enable them to achieve a high score on a posttest. In order to examine the score gain, it is recommended that the students' pretest scores be compared with their posttest scores. This procedure is termed intervention strategy (Brown, 2005). Study of the score gains can serve as empirical support showing that learning has taken place. Conversely, if gains are not observed for certain objectives, teachers should reconsider their teaching plans to better enable effective learning.

Five different kinds of software were used in this study. Apart from GENOVA, the other four software programs are quite user-friendly. The teachers can refer to the output to confirm whether their experience-derived teaching is suitable for the actual outcome of the teaching. Although this requires hard work, it is definitely beneficial in terms of improving teaching.

Two limitations are inherent in this study. First, further investigation is required to determine which criterion-referenced, multiple-choice vocabu-

lary test items are valid. Second, replication studies should be conducted to investigate how the magnitude of the G coefficient for absolute decisions in criterion-reference language tests would change in different testing situations. In spite of the fact that CRTs are frequently used by many teachers, studies on CRTs are rarely conducted. Additional studies on this issue are needed of other language programs in Japanese university contexts to reveal ways to prepare dependable and valid CRTs.

Acknowledgement

I would like to thank Dr. J. D. Brown for his invaluable lessons on language testing. I am also grateful to H. P. L. Molloy, two anonymous reviewers, and the editor of *JALT Journal* for reading earlier drafts of this paper and providing me with insightful comments. I am solely responsible for any remaining errors.

References

Assessment Systems Corporation (1995). *XCalibre* (version 1.1e) [computer software]. St. Paul, MN: Assessment Systems Corporation.

Assessment Systems Corporation (1996). *ITEMAN* (version 3.6) [computer software]. St. Paul, MN: Assessment Systems Corporation.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. (1997). Generalizability theory. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (pp. 255-262). Norwell, MA: Kluwer Academic Publishers.

Bachman, L. F. (2004). *Statistical analysis for language assessment*. Cambridge: Cambridge University Press.

Brennan, R. L. (1980). Applications of generalizability theory. In R. A. Berk (Ed.), *Criterion-referenced measurement: The state of art* (pp. 186-232). Baltimore, MD: Johns Hopkins University Press.

Brennan, R. L. (1983). *Elements of generalizability theory*. Iowa City, IA: American College Testing Program.

Brennan, R. L. (2000). Performance assessments from the perspective of generalizability theory. *Applied Psychological Measurement, 24*(4), 339-354.

Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag.

Brennan, R. L., Gao, X., & Colton, D. A. (1995). Generalizability analysis of work keys listening and writing tests. *Educational and Psychological Measurement, 55*(2), 157-176.

Brown, J. D. (1982). *Testing EFL reading comprehension in engineering English*. Unpublished doctoral dissertation, University of California, Los Angeles.

Brown, J. D. (1993). A comprehensive criterion-referenced testing project. In D. Douglas, & C. Chapelle (Eds.), *A new decade of language testing: Collaboration and cooperation* (pp. 163-184). Ann Arbor, MI: University of Michigan.

Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.

Brown, J. D. (1999). The relative importance of persons, items, subtests and languages to TEOFL test variance. *Language Testing, 16*(2), 217-238.

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. New York: McGraw-Hill.

Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.

Brown, J. D., & Ross, J. A. (1996). Decision dependability of subtests, tests and the overall TOEFL test battery. In M. Milanovic, & N. Saville (Eds.), *Performance testing, cognition and assessment* (pp. 231-265). Cambridge: Cambridge University Press.

Crick, J. E., & Brennan, R. L. (1983). *GENOVA* (version 2.1) [computer software]. Iowa City, IA: The American College Testing Program.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology, 16*, 137–163.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: Wiley.

Cronbach, L. J., Linn, R. L., Brennan, R. L., & Haertel, E. H. (1997). Generalizability analysis for performance assessments of student achievement or school effectiveness. *Educational and Psychological Measurement, 57*(3), 373-399.

Denies, J. (2004). *WordClassifier* (version 2.5). [computer software]. Belgium: Michael Goethals and EET Project Team.

Fyans, J. L. (Ed.). (1983). *Generalizability theory: Inferences and practical application*. San Francisco: Jossey-Bass Inc.

Griffee, D. (1995). Criterion-referenced test construction and evaluation. In J. D. Brown & S. Yamashita (Eds.), *Language testing in Japan* (pp. 20-28). Tokyo: The Japan Association for Language Teaching.

Kunnan, A. J. (1992). An investigation of a criterion-referenced test using G-theory, and factor and cluster analyses. *Language Testing*, *9*(1), 30-49.

Lynch, B. K., & McNamara, T. F. (1998). Using G theory and many facet Rasch measurement in the development of performance assessment of the ESL speaking skills of immigrants. *Language Testing, 15*(2), 158-180.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.

Strube, M. J. (2000). Reliability and Generalizability theory. In L. G. Grim, & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 23-66). Washington, DC: American Psychological Association.

Suen, H. K. (1990). *Principles of test theories*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Thompson, B. (2003). A brief introduction to Generalizability theory. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 43-58). Thousand Oaks, CA: Sage.

Yamanishi, H. (2004). How are high school students' free compositions evaluated by teachers and teacher candidates: A comparative analysis between analytic and holistic rating scales. *JALT Journal, 26*(2), 187-205.

Endnotes

1. $X = T + E$
   where: X = observed score
   T = true score
   E = error

2. $r_{xx} = \sigma^2(t) / (\sigma^2(t) + \sigma^2(e))$
   where: r_{xx} = reliability
   $\sigma^2(t)$ = true score variance
   $(\sigma^2(t) + \sigma^2(e))$ = observed score variance

3. $\alpha = (k/k{-}1)\ (1{-}\sum\sigma^2(i)/(\sigma^2(t) + \sigma^2(e))$
   where: α = Cronbach alpha internal consistency reliability
   k = number of items
   $\sum\sigma^2(i)$ = sum of items variance
   $(\sigma^2(t) + \sigma^2(e))$ = observed score variance

4. $SEM = SD_x\sqrt{}\ (1 - r_{xx})$
   where: SEM = standard error of measurement
   SD_x = standard deviation of the test score
   r_{xx} = reliability

5. $r_{kk} = kr_{xx}/(1 + (k - 1)\ r_{xx})$
   where: r_{kk} = estimated reliability when the multiple of test items is set at k
   k = number of items
   r_{xx} = reliability

6. $E\rho^2(\delta) = \sigma^2(p)/(\sigma^2(p) + \sigma^2(pi)/ni)$
   where: $E\rho^2(\delta)$ = G coefficient for relative decisions
   $\sigma^2(p)$ = persons variance
   $\sigma^2(pi)$ = persons-by-items interaction
   ni = number of items

7. $E\rho^2(\Delta) = \sigma^2(p)/(\sigma^2(p) + (\sigma^2(i)/ni) + (\sigma^2(pi)/ni))$
   where: $E\rho^2(\Delta)$ = G coefficient for absolute decisions
   $\sigma^2(p)$ = persons variance
   $\sigma^2(i)$ = items variance
   $\sigma^2(pi)$ = persons-by-items interaction
   ni = number of items

タスクの複雑性、母語、日本語能力が談話構成に及ぼす影響—指示表現・受身表現の使用について—

Effects of Task Complexity, L1 and Proficiency on the Use of Referential and Passive Forms in L2 Japanese Discourse

中浜 優子（なかはま　ゆうこ）
慶應義塾大学

This study investigates how task complexity, learners' L1 backgrounds, and proficiency levels influence the ways in which learners introduce and maintain referential topics in L2 Japanese oral narratives. Narrative discourse produced by two groups of learners of Japanese, one whose L1 is typologically parallel to and one whose L1 is distant from Japanese (Korean and English, respectively) was investigated at two different proficiency levels (intermediate and advanced levels, based on the ACTFL guidelines). Native speakers (NS, hereafter) and nonnative speakers (NNS, hereafter) of Japanese constructed two narratives: looking at a series of pictures, and reconstructing a story after seeing a silent film. The former task was supposed to elicit narratives in a "here and now" context (H/N, hereafter), while the latter, in a "there and then" context (T/T, hereafter). The latter task is considered more complex than the former, as in the T/T task, the speakers would need to code, store, and search the content of the story in viewing the film as well as retelling the story, and thus it is cognitively more challenging than the H/N setting.

The results showed that different narrative tasks brought out differing topic management patterns in both NS and NNS narratives. Especially notable was that the task complexity affected the L2 Japanese narrative discourse produced by lower proficiency level English speakers. The T/T narrative task revealed higher grammati-

JALT Journal, Vol. 31, No. 1, May, 2009

cal accuracy in English speakers' narratives, such as the use of NP+*ga*' and non-use of particle omissions. Yet, English speakers under-produced zero anaphora in their T/T narratives, whilst the production of zero anaphora reached target-level in their narratives in the H/N context. In contrast, Korean speakers showed consistent grammatical accuracy irrespective of task types or proficiency. It was argued that positive L1 influence was observed in Korean speakers' Japanese L2 narratives in referential topic management, and negative L1 influence was observed in English speakers' corresponding narratives, as Korean and Japanese share comparable linguistic features with respect to code topic continuity and discontinuity. While Korean speakers showed similar patterns of use of particles and zero anaphora to NS of Japanese throughout the narratives, their performance in the production of passive structures differed from NSs, and resembled the performance of English speakers. As proficiency increased, both learner groups produced more passive structures. The advanced level learners produced nearly twice as many passive structures in H/N than T/T, whereas task complexity did not have an influence on the number of passive structures produced in NS narrative discourse. Moreover, even advanced level learners in the H/N context produced fewer instances of passive structures than their NS counterparts. It is argued that this resulted from the different ways in which Japanese and English/Korean speakers place their focus in developing a story in their L1. Specifically, Japanese tend to place their focus on the main characters and tell a story from their viewpoints, whereas English and Korean speakers put their focus on action or fact. Japanese speakers in the current study kept main characters in the topic position; sometimes as agent in active structure, other times as patient in the passive structure. English and Korean speakers tended to switch the topic of the sentence between protagonists and antagonists, instead of keeping the protagonists in the topic position, and as a result, produced fewer instances of passive morphemes.

本研究は、学習者の母語、日本語能力、タスクの複雑性が、ナラティブ談話のトピック管理に及ぼす影響を調査したものである。韓国語・英語を母語とする中・上級レベルの日本語学習者と日本語母語話者に、目の前にあるものを描写するナラティブ（同時空間：H/N）と、記憶に頼りながら描写するナラティブ（非同時空間：T/T）を口頭で語ってもらった。

韓国語話者は英語話者に比べ、助詞とゼロ照応の適切な使用等、トピック管理能力が早くから備わっており、タスク間での差はなかった。英語話者においては、T/Tタスクの方がH/Nより助詞の正用を引き出す傾向にあったが、ゼロ照応の産出に関してはH/Nの方が、産出割合が増した。受身表現に関しては、学習者両グループとも、日本語母語話者より産出数が少なかった。これは、立場志向を取る日本語談話パターンと事実志向の英語・韓国語談話パターンの違いが受身形の産出数の違いという形で現れたものだと思われる。

Keywords: Japanese oral narrative, task complexity, referential topic

はじめに

文脈内での言語使用を重視する機能主義言語学の考えでは、指示対象トピックの適切な導入・維持が、物語談話の一貫性向上につながる一要因だと捉えられている。談話におけるトピックの連続性の高さを示すマーカーとして、代名詞やゼロ照応が挙げられるが、日本語や韓国語などの言語では、トピックの連続性の高い名詞を照応する際には、照応詞の省略、即ち、ゼロ照応の使用が頻繁に見受けられる。[1]

第二言語（L２）習得分野の先行研究（Nakahama, 2003a, Polio 1995, Yanagimachi 2000等）では、ゼロ照応の適切な使用は、学習者の日本語能力と共に増えることが明らかになっている。また、英語と日本語ではトピックの連続性が比較的高いとされる定性マーカー（英語では定冠詞、日本語では「は」）の方が連続性の低い不定性マーカー（英語では不定冠詞、日本語では「が」）より習得が早いことが分かっている（土井・吉岡1990, Chaudron & Parker 1990, Clancy 1985, Huebner 1985, Master, 1997, Sakamoto 1993等）。しかし、Nakahama（2003b）では日本語学習者にとって困難であろうとされる「が」においても、日本語に近似した助詞のシステムを持つ韓国語を母語とする学習者にとっては、初中級レベルで既に日本語母語話者（NS）と同様、適切な使用がなされていることが報告されている。

また、タスクの複雑性が物語談話における文法の正確性、談話の複雑性及び流暢性に与える影響ついても指摘されている（Robinson, 1995）ことから、本研究では、学習者の母語、日本語運用能力とタスクの複雑性が、L２物語談話の導入・展開におけるトピック連続性指標の選択（具体的には指示表現と受身表現の使用）に及ぼす影響を明らかにすることを目的とする。

先行研究

談話内でのトピック管理

Givón（1983）は、談話というものは、多命題で構成されており、いくつかの節が統合されて成り立つテーマ段落として捉えるべきだと主張し、談話の一貫性を左右する重要な要因として、トピックの連続性・非連続性を挙げている。そのGivónの異言語間トピック連続性の尺度によると、談話内の焦点となっている最も連続性の高いトピックは、ゼロ照応で記され、新情報として談話に導入された、最も非連続的なトピックは、完全な形の名詞句で記される。すなわち、物語談話において、指示対象物を最初に導入する際は、完全名詞句で表し、話題のトピックとして維持され、連続して言及される場合は、ゼロ照応として表されるというのである。Givón（1983）やTomlin（1990）などの機能言語主義者の考えでは、トピックの導入・維持、すなわち話者が談話内でトピックをいかに適切に管理して行くかによって、談話の一貫性が左右されるというのである。

英語では、新情報導入には、名詞句＋不定冠詞を使用し、一度導入されたもの、すなわち旧情報について言及する場合には、名詞句＋定冠詞を使用するのがプロトタイプであるとされている。また、話題のトピックとして連続して言及する場合は、代名詞句が使用される（Chaudron & Parker, 1990）。それに対し、日本語では、主語/トピックの位置で新しい指示対象について言及するには、助詞「が」を用い、既知の指示対象には助詞「は」を用いる。最も高いトピック連続性を示すためには、ゼロ照応が用いられる（Hinds 1983）。韓国語でのトピック導入・維持のパターンは日本語と類似し

ており、新しい指示対象を導入する際は「가/이」を、既知の指示対象には「는/은」を用い、トピック連続性が高い対象物には、ゼロ照応が使用される（Kim 1989）。

談話内での名詞句・照応詞習得における言語転移

Li & Thompson's（1976）の主題・主語卓越類型モデルを援用した研究の中には名詞句・照応詞習得における第一言語（L1）からの言語転移について考察されたものがあるが（Fuller & Gundel 1987, Gundel & Tarone 1983, Huebner 1983, Jin 1994等）、未だ統一した見解は得られていない。Huebner（1983）やRutherford（1983）が言語習得の初期の段階において見られる主題卓越性を、母語からの言語転移と主張するのに対し、Fuller & Gundel（1987） では、ゼロ照応の出現は、初期に見られる主題卓越の段階の一つの普遍的特徴であると提言している。しかしJin（1994）やPolio（1995）が行った主題卓越言語の習得研究では、Fuller & Gundelが主張した、習得の初期に見られるという主題卓越性は見られなかったことから、主題卓越の普遍説は覆されることとなった。Polio（1995）、Nakahama（2003b）では、それぞれ中国語習得、日本語習得において、母語に主題卓越言語をもつ学習者にとっても、初級レベルでは、ゼロ照応の使用は困難であったことが報告されている。

一人称・二人称・三人称の３種のナラティブを比較分析したYanagimachi（2000）の研究では、最初から視点の移動の余地がない一人称・二人称のナラティブの場合は、ゼロ照応の運用に問題はなかったものの、自ら視点を決めて話をしていくというタスクになれば、学習者と母語話者間の談話構成に大きな違いがでることが分かった。

このYanagimachiの研究結果からも分かるように、視点とトピック管理の密接な関連は明白である。日本語が立場志向の傾向を持つ言語であるのに対し、英語、韓国語、中国語などは事実志向の言語であると捉えられており（日英比較に関する詳しい考察は水谷1985参照）、日本語を母語とする話者が、自ら視点を特定の登場人物にあて、その人物を中心に話を構築していくのに対し、英語などの母語話者の場合、起こった事象に焦点をあてる。日本語母語話者では、特定人物に視点をおくため、中心人物をトピック性の高い位置（主語の位置）においたまま、授受表現や受身形の使用などにより、ストーリーを展開していく傾向が見られるが、事実志向の言語を母語に持つ学習者が日本語を話した際、授受表現や受身形の使用数は、上級レベルに至っても、母語話者に比べ大幅に下回る傾向があることが報告されている（金2001, Nakahama 2003b, 田代 1995, 渡辺1996, Yanagimachi 2000等）。また、他の文法項目に比べても、視点や受身形は、学習者にとって習得がより困難であることも分かっている（田中, 1996, 1997）。これらの結果から、視点のおき方という概念・意味の言語的表出に関しては、習得が難しいというだけではなく、母語からの影響が避けがたいということが分かる。（意味の転移、言語の相対性についての詳しい考察はOdlin, 2002, 2005等を参照のこと）

Nakahama（2003b）では母語からの言語転移を調査するため、韓国語と英語を母語とする学習者グループ（中級下・中級上・上級）を比較した。その結果、韓国語話者の発話で、「が」「は」などの助詞の使用に母語からの正の転移が見受けられたものの、ゼロ照応の使用においては、中級下レベルでは学習者の言語的制約が正の転移の妨げになったことが分かった。また、英語話者に関しては、日本語運用能力が低いと、「が」の習得が難しく、「は」で代用されていることが分かり、定性（日本語では「は」、英語では定冠詞（the））の習得が不定性（日本語では「が」、英語では不定冠詞（a, an））の習得より先行すると示されている先行研究（日本語習得はClancy 1985,

Sakamoto 1993等、英語習得はChaudron & Parker 1990, Huebner 1985, Master, 1997等参照のこと）や、定性を表す言語形式の使用がコンテクストに関わらず過剰般化され、その正用が熟達度に比例し増えていくという研究結果（Andersen 1977, 遠山2005, Huebner 1983等参照）を支持することとなった。

目の前にある物を描写しながらストーリーを構築する方法（Here and Now, 以下H/N）でデータ収集を行ったNakahama（2003b）では、最もトピック連続性の高い文脈において、使用された全言語形式を分母とした際のゼロ照応の使用割合は、英語を母語とする日本語中級下レベルで61.3%、中級上レベルでは84.8%であった。それに対し、無声映画を見た後でストーリーを思い出しながら再構築する方法（There and Then, 以下T/T）の抽出法を用いたNakahama（2003a）の英語母語話者の日本語中級レベルではゼロ照応使用率が56%であったことから、ゼロ照応の使用率にタスクの種類（ナラティブ抽出法）が何らかの影響を及ぼした可能性が示唆される。

Robinson（2003）によると、複雑なタスクは簡単なタスクよりも注意の負荷が高く、この負荷の違いが学習者の発話の正確性、流暢性、複雑性に影響を及ぼすとしている。T/TとH/Nのタスクを比べてみると、前者では、登場人物等、物語の内容を認知的手段により貯蔵、構築しなければならないことから、H/Nに比べ認知面で負担が増えることから、より複雑なタスクであると言える（Ishikawa, 2005）。この認知面での負担の高低が、二つのタスクでのトピック連続性の高い文脈におけるゼロ照応と完全名詞句の産出割合に影響を及ぼしたのではないかと考えられる。

H/NとT/Tのタスクにより産出された英語のL2ナラティブに見られる文法の正確性、複雑性、流暢性などを取り扱った研究にRobinson（1995）がある。Robinsonの研究によると、タスクとしての複雑性の高いT/Tの方が、抽出されたナラティブの文法の正確性・複雑性が高いということが分かった。文法の正確性の一つとして、冠詞の使用について考察しているが、T/Tの方が、H/Nよりも冠詞の正用が多いことが報告されている。Robinson（1995）はT/Tに見られる冠詞の正用率の高さについて、単純に絵を描写していくH/Nに比べ、話者が、命題間をうまく繋いで、結束性のあるストーリーを構築するため、多大な注意を払っていった結果、T/Tのタスクでより高い正確性が得られたのではないかとしている。後に、Robinson（2004）は認知仮説（Cognition Hypothesis）として、複雑性の高いタスクは、学習者が備え持つ言語的資源の活用をより多く引き出す、と唱えている。以上、概観してきた様々な先行研究の結果を踏まえ、本研究では、以下の2つの研究設問を明らかにすることを試みる。

研究設問

1）日本語と類型論的に近似した言語を母語とする学習者と相違した言語を母語にもつ学習者間では、物語談話におけるトピック管理能力の発達に差異が見られるか。
2）タスクの複雑性は、話者のトピック管理にどのような影響を及ぼすのか。また、タスクの複雑性が与える影響は、学習者と母語話者で類似するのか。

方法

被験者

本研究の被験者は、日本で日本語を学習している、英語を母語とする中級レベル（EM）11名、上級レベル（EH）10名、韓国語を母語とする中級レベル（KM）10名、

上級レベル（KH）10名、及び日本語母語話者（NS）10名であった。

日本語非母語話者の被験者のレベル分けであるが、本研究のタスクを行う前に、Simulated Oral Proficiency Interview（SOPI）のタスクの一部を行ってもらい[2]、ACTFLガイドラインに基づき、筆者が学習者の日本語能力を判定した。中級中、中級上と判定された学習者を中級グループに、上級下、中、上と判定された学習者を上級グループに振り分けた。それぞれのレベルから三つずつSOPIタスクのデータを無作為抽出し、別のSOPI評価者にレベル判定してもらい、評価者間の信頼性を確かめたところ、Cohenのカッパ係数は0.86で、評価者間の一致度が高いことが分かった。

被験者の平均年齢はNS（20.3歳）、EM（21.5歳）、EH（26.8歳）、KM（23.0歳）KH（24.2歳）、男女比はNS（男:5、女:5）、EM（6:5）、EH（6:4）、KM（6:4）KH（3:7）であった。日本語学習歴、日本滞在年月の平均はそれぞれ、EM（50.7ヶ月、9.5ヶ月）、EH（64.7ヶ月、32ヶ月）、KM（18.9ヶ月、9.5ヶ月）、KH（57.3ヶ月、18.3ヶ月）であった。

タスク　（H/NタスクとT/Tタスク）

前述のH/N、すなわち、目の前にある物を描写する、同時空間でのストーリー構築データ抽出のため、文字のない24枚の連続した絵からなる絵本Frog, *Where Are You?*（Mayer 1969）を使用し、それを見ながらストーリー構築をしてもらった。また、非同時空間でのストーリー構築、すなわちT/Tタスクとしては、無声映画*Winter Carousel (from The Cameraman's Revenge and Other Fantastic Tales* by Image Entertainment）をまず観賞した後、話を頭の中でまとめて、映画を見ずに、その内容を被験者に口頭で語ってもらった。[3]　両タスクとも発話を３０分以内で終了するように指示を出した。どちらの物語も、３名（匹）の主人公がおり、それらを取り巻く脇役的存在の登場人物（動物）が出てきてストーリーが展開されていくというものである。Robinson（2001）等の先行研究で、タスクの順序が言語産出に影響を及ぼす可能性が示唆されていることを踏まえ、影響を最小限にするため、全てのグループの半分の被験者はH/N—>T/Tの順に、後の半分はその反対の順序でストーリー構築をしてもらった。

ナラティブの聞き手は、筆者と研究補佐２名が担当した。聞き手は、質問は控え、発話は相づち程度に抑え、話し手には基本的に独話調で話してもらった。

フォローアップ面接

フォローアップ面接をNS全員と学習者数名（EM３名、KM２名、EH３名、KH３名）に実施した。インタビューでは、自分の発話の内容をテープで聴いてもらい、ストーリーの中でのトピック管理のための「が」、「は」などの助詞、ゼロ照応、受身形の使用などについて調査した。

分析

本稿では、「トピック管理」の操作上の定義を「トピックとなり得る指示対象の物語談話への導入と維持」とし、日本語学習者、母語話者の物語談話におけるトピック管理能力について詳しく分析する。それに際し、同じ絵本でデータ収集をしたNakamura（1993）、Nakahama（2003b）の枠組みを援用し、1）指示対象の導入、2）

トピックスイッチ、3）トピックの連続的な言及の3つの文脈ごとに、データを文字化・コード化し分析した。指示対象の導入とは、「任意の文要素で談話に最初に導入されたもの」であり、トピックスイッチとは、「既に物語に導入されている指示対象で、（トピック性を失っていたものが）談話のトピックとして主語の位置で再導入されたもの」と定義する。トピックの連続的な言及とは「主語の位置におけるトピックスイッチの直後に（主語の位置で）産出され続ける指示対象」とする。指示対象導入の文脈は、トピックの連続性が最も低く、トピックの連続的言及の文脈では、トピックの連続性が最も高いと言える。以下、1）に指示対象導入の例を、2）にトピックスイッチとトピックの連続的な言及の例を示す。

1）男の子が蛙を捕まえてきてそれを瓶に入れました。

この例では、男の子の導入には「が」が使用され、蛙の導入には「を」が使用されている。

2）犬が窓から飛び降りてしまいました。男の子は怒っています。（∅）森の中へ行って（∅）穴を探しました。

この発話文は、ナラティブ開始の数発話後に出てきたものである。ここに至るまでの段階で、男の子と犬の導入はすでに終わっており、ここでは、犬が談話のトピックとして再導入されて始まっている。しかし、談話のトピックは男の子に移行し、名詞と「は」によって主語の位置に現れてきている。この「が」と「は」双方とも、トピックスイッチを表す言語形式とみなされる。その後、男の子がトピックのままで、話が続いていくが、「森の中へ行く」、「穴を探す」という動作の主語として照応詞は使われず、ゼロ照応で示されているのが分かる。ここで見られる二度のゼロ照応（∅）の使用にあたる部分が、トピックの連続的な言及の例である。

コード化の客観性を保証するため、学習者グループからデータを3つずつ（計12データ）無作為に選び、別の日本語母語話者にもコード化してもらい、結果を比較したところ、Pearsonの相関係数は.995で、高い相関関係があることが分かった。

データの分析方法は、それぞれの文脈で使用された名詞句の出現頻度ではなく、その使用割合の平均値を比較した。[4]　これは、被験者によって発話の長さが統一されておらず、それゆえ合計回数を従属変数とした場合、発話数が短い被験者のナラティブでは、ある言語形式の出現頻度は、発話数の長い被験者のナラティブより必然的に低くなる可能性が高い。つまり、公平な比較ができなくなる。それに対し、平均使用割合を比較すれば妥当性の高い検証に導くことができる。ただ、助詞脱落など、産出率の非常に低い文法項目もあった。こういったケースでは分散分析を行う上での分布規定の条件が犯されることになるが、関連の模擬実験研究（Harwell, Rubinstein, Hayes, & Olds 1992等参照のこと）では、分散分析での分布規定が犯されても、その統計処理による結果の妥当性は変わらないという結果が出ていることから、本研究では統計処理は分散分析で行う。

まず、各文脈（指示対象導入・トピックスイッチ・連続的言及）での使用言語形式を調べ、また受身形の出現数についても結果を報告する。

結果と考察

指示対象の導入

日本語母語話者と日本語学習者が産出した、H/NとT/Tでの指示対象の導入の割合の平均値をそれぞれ表1の上段、下段に示す。頻繁に現れた言語形式、及び各グループに特徴的な項目以外は「その他」としてまとめて記載する。

日本語学習者の産出データの指示対象導入の分析において、韓国語話者・英語話者両グループで、「が」「は」「NP＋φ（助詞の脱落）」の使用に特徴的パターンが見られた。以下にその特徴を明らかにする。

1.「NP+が」と「NP+は」の使用

指示対象を導入する際、H/N、T/T両タスクにおいて、NSだけでなく、学習者も日本語レベルや母語に関わらず、「が」の使用割合が最も多いことが分かった。5グループの平均は、H/Nで48.1%、T/Tで60.8%であった。タスク間の被験者内反復測定を行った結果、タスク間の差は有意であった（F（1,46）=11.8, p=.002,　偏η^2=.198）。

母語別、タスク別に見てみると、英語話者の場合、H/Nでの「が」の産出割合は、中級レベルでは低いが（28.5%）、上級になると増加し（47.6%）、NS（52.9%）に近づいて行っているのが分かる。一方、韓国語話者の場合、レベル間でそれほど違いが見られず、むしろ、中級の方が、「が」の産出率が高かった（中級：63.1%、上級：50.5%）。NS、学習者の総5グループを比較し、分散分析を行ったところ、「が」の使用割合にグループ間での有意差が見られた（F（4,46）=3.81,　p=.009,偏η^2=.249）シェフェーの多重比較により、有意差はEMとKMにあることが分かった（p＜.05）。T/Tでは、グループ間で、「が」の使用率に有意差は見られなかった。

表1 指示対象導入に使われた言語形式: H/N、T/Tタスク
（数値は使用形式ごとの割合の平均値）

NP+	が	は	を	に	も	助詞脱落	その他
英語中級							
H/N	28.5	12.3	13.8	5.8	9.5	16.5	13.5
T/T	41.5	31.0	3.6	6.7	7.1	0	10.1
韓国語中級							
H/N	63.1	0	14.6	2.0	2.9	0	17.5
T/T	71.8	3.1	5.0	4.2	6.4	0	9.5
英語上級							
H/N	47.6	10.9	11.0	10.0	2.5	0	17.9
T/T	64.4	2.0	1.7	6.4	13.0	0	12.6
韓国語上級							
H/N	50.5	4.5	12	7.8	2	0	21.2
T/T	65.2	1.7	0	6.7	4.5	5.0	17.0
母語話者							
H/N	52.9	0	19.7	11.1	4.5	0	13.3
T/T	63.0	0	0	7.4	5.2	0	24.5
合計平均値							
H/N	48.1	5.7	14.1	7.3	4.4	3.6	16.5
T/T	60.8	8.0	2.1	6.3	7.2	1.0	14.7

注　H/NはHere and Now, T/TはThere and Thenのタスクを表す。

定性マーカー「は」の使用については、NSでは見られず、韓国語話者ではタスクタイプ、日本語熟達度に関係なく、産出率は低かった。一方、英語話者ではH/Nでは中級で12.3%、上級ではやや減少し10.9%であった。

前述のように、T/Tでは指示対象を導入する際、全体的に「が」の使用が多く見られたが、EMグループでもその傾向があり、H/Nの28.5%から41.5%まで使用率が上がっている。その結果、「は」の使用率も減少するのが予測されたのであるが、12.3%から31.0%まで使用率が上昇していた。EMグループは、H/Nでは「が」「は」以外に助詞の脱落、「を」の使用などの言語形式も用いられていたのに対し、T/Tでは、助詞の脱落は見られず、「を」の使用も3.6%にとどまっていた。また（ここでは「その他」の中に含まれる）助詞の誤りがH/Nで6.1%ではあるが見られたのに対し、T/Tでは全く見られなかったという違いもあった。

英語話者グループのT/Tでの「は」の使用割合は、上級レベルでは2%と使用率が中級レベルに比べると大幅に下がっており、これは、上級レベルでは「が」の使用率が大幅に増えたためだと思われる。母語話者・非母語話者5グループを比較したが、EMグループにおける「は」の平均使用割合は他の４グループに比べて有意に高いのが分かった（F（4,46）=7.82,　p=.000, 偏η^2=.405）。このことから、H/N同様、英語話者の「は」の適切な使用（不使用）に関しては、日本語能力と正比例し、韓国語話者に関しては、中級レベルから既にNSと類似したパターンを提示しているのが分かった。

以上、「が」と「は」の使用（不使用）についてタスク間、グループ間比較をしてきたが、何故T/Tでの方が、「が」の産出割合が増え、また、文法の正確性も上がったのだろうか。

Robinson（1995）でも、T/Tで物語を構築する時は、H/Nよりも、英語における冠詞の正用が増えるという報告があったが、これは、話者が頭の中に貯蔵してある内容を慎重にアクセスし、まとめながら話すため、新しい登場人物を導入する際、それだけ文構造にも「注意」を払うようになり、文法の正確性が上がったのではないかと考えられる。T/T　における、EMの文法の正確性は「が」の使用率の増加だけではなく、助詞の脱落、助詞の誤りがなくなったことからも垣間見られた。H/Nでは、前者が16.5%、後者が6.1%見られたのに対し、T/Tではどちらも0%であった。Robinsonの英語習得研究と本研究の日本語習得研究の結果から、タスクの複雑性がナラティブ構築における文法の正確性に影響を及ぼす可能性は、普遍的なものではないかと考えられる。

T/Tでは学習者両グループにおいてレベル間の差はなかったのであるが、H/Nでは英語話者の場合、「が」の使用率は日本語熟達度に比例して上昇しており、L2習得において、不定性マーカーの正用が外国語能力レベルとともに増えていくという先行研究をも支持する結果となっている。一方、日本語と同じ後置詞のシステムを持ち、「が」と「は」に対応する助詞を使用する韓国語話者においては、タスクタイプや日本語能力に関係なく、母語話者と同じように首尾一貫した「が」の適切な使用が見られた。

全てのNSと学習者の一部にフォローアップ面接を行ったのであるが、NSは「が」と「は」の使用法について、意識的に使用してはいないことが分かった。学習者では、インタビューをしたEM３名とも、「が」と「は」を差別化しておらず、どちらも主語を示す助詞だという回答が得られ、EH３名のうち２名は、「が」と「は」の用法について言語学的知識を持っているのが分かった。韓国語話者においては、中級、上級５名の

学習者全員が、韓国語にも「が」と「は」に対応する後置詞があることから、「が」、「は」に韓国語の助詞「가、이」、「는、은」を対応させて使用していることが分かった。このことから、母語と目標言語の距離が近い場合、両言語の文法的類似点を把握しており、それをうまく活用し、第二言語としての日本語の談話構成に役立てていることが分かる。それに対し、英語話者のように、母語に対応するものがない言語項目を習得する場合、指標にするものが存在しないため、長年の学習と肯定的・否定的証拠にさらされることにより、上級レベルまで進んで始めて、談話においての適切な「が」と「は」の使い分けができてくるのではないかと思われる。[5]

2. 「NP+助詞の脱落」[6]

助詞の脱落がH/NのEMグループで16.5%も見られ、「が」の28.5%の次に使用率が高いのが分かった。これは先行研究（Nakahama 2003b）の結果を支持するものとなり、H/Nでは、英語を母語とする日本語学習者は上級に達するまでは、助詞を脱落させる傾向がある事が改めて確認された。英語には助詞のシステムがないため、負の転移が助詞の脱落という形で現れた可能性が高い。

助詞が脱落された箇所を見てみると、本来「を」や「が」を使用すべき箇所での脱落であり、また、助詞が脱落された名詞句は主人公であったり脇役であったりしたことから、EMはランダムに助詞を脱落させていることも分かった。H/Nでは、NS、韓国語話者に関しては、助詞の脱落は見られなかった。

T/Tにおいては、助詞の脱落は、KHの5%の産出という例外はあったものの、EMを含む4グループで検出されなかった。これは、前述の、非同時空間における事物言及の際の文法の正確性に貢献するものであり、Robinson（1995）の研究結果を支持することとなった。

以上、指示対象導入の際に使用された主要言語形式について、タスク・グループ間比較をしてきた。T/Tでの全グループにおける「が」の使用割合増加、EMにおける文法の正確性向上などが確認され、同時空間・非同時空間でのストーリー構築という、タスクの複雑性の違いが助詞の産出に影響を及ぼした可能性が指摘できる。

トピックスイッチ（指示対象の再導入）

トピックスイッチの文脈では、「が」、「は」、「も」、助詞の脱落、ゼロ照応の5つの言語形式の使用が見られた。両タスクでの言語形式の割合の平均値を表2に示す。

トピックスイッチの文脈において使用率の高かったものに、「が」と「は」が挙げられる。各グループの使用率を比較したところ、その使用割合に有意な差はなかった（H/Nにおける「が」の使用：$F(4,46)=0.73$, $p=.58$,偏$\eta^2=.059$、「は」の使用：$F(4,46)=0.63$, $p=.65$,偏$\eta^2=.052$、T/Tにおける「が」の使用：$F(4,46)=0.54$, $p=.71$,偏$\eta^2=.045$、「は」の使用：$F(4,46)=1.1$, $p=.37$,偏$\eta^2=.09$）。先行研究では、助詞別に分けず、完全名詞句として、ゼロ照応との使用割合を比べている場合が多い（Clancy 1985, Yanagimachi 2000等）ので、本研究でも、「が」、「は」、「も」の合計を完全名詞句とし、ゼロ照応の使用頻度と比較することにする。

グループ間比較の際は話者によって発話量が違うことから、公平性を考慮に入れ、平均使用割合を比較してきたが、今回は被験者間ではなく被験者内での項目比

表2 トピックスイッチに使われた言語形式: H/N、T/Tタスク
（数値は使用形式ごとの割合の平均値）

NP+	が	は	も	助詞脱落	ゼロ照応
英語中級					
H/N	36.3	30.1	0.6	11.2	21.8
T/T	36.7	42.4	2.2	11.9	6.7
韓国語中級					
H/N	33.5	43.4	1.6	0	21.5
T/T	45.3	31.9	0.9	4.2	17.8
英語上級					
H/N	41.2	36.9	3.6	0.4	18.0
T/T	52.1	25.5	7.7	4.1	10.6
韓国語上級					
H/N	29.7	44.8	1.8	0	23.7
T/T	46.7	23.4	1.9	6.0	22.0
母語話者					
H/N	25.1	33.1	2.3	0	39.6
T/T	40.0	30.5	2.9	2.1	24.1
合計平均値					
H/N	33.2	37.5	1.9	2.5	24.8
T/T	44.1	31.0	3.1	5.8	16.0

注　H/NはHere and Now, T/TはThere and Thenのタスクを表す。

較なので、使用頻度の合計をカイ二乗検定を用いて検定した。その結果、両タスクとも全てのグループで、完全名詞句の使用頻度がゼロ照応のそれよりも1％水準で、有意に多いことが確認できた。[7]

各グループでトピックスイッチの対象となった登場人物を、主要人物（主人公）と非主要人物（脇役）に分類したものを表3に示す。

表3にあるように、トピックスイッチの対象となった登場人物は、両タスクで大多数（H/N：92.5％、T/T：89.8％）が主要登場人物であることが分かり、先行研究（Clancy, 1992）を支持する結果となった。

ゼロ照応

トピックスイッチにおけるゼロ照応の使用割合の総グループ平均はH/Nで24.8％、T/Tで16.0％であった。被験者内タスク間の比較を検定した結果、トピックスイッチの文脈

表3　トピックスイッチの対象人物
（一段目の数字は使用頻度、二段目のカッコ内の数字は使用割合）

	H/N　主人公	H/N　脇役	T/T　主人公	T/T　脇役
英語中級	100 (91.7%)	9 (8.3%)	117 (88.0%)	16 (12.0%)
韓国語中級	173 (94.0%)	11 (6.0%)	113 (94.2%)	7 (5.8%)
英語上級	239 (89.2%)	29 (10.8%)	182 (87.5%)	26 (12.5%)
韓国語上級	205 (94.5%)	12 (5.5%)	156 (90.2%)	17 (9.8%)
母語話者	105 (94.6%)	6 (5.4%)	143 (90.5%)	15 (9.5%)
合計	822 (92.5%)	67 (7.5%)	711 (89.8%)	81 (10.2%)

注　H/NはHere and Now, T/TはThere and Thenのタスクを表す

で使用されたゼロ照応の使用割合はタスク間で有意に違うことが分かった（F（1,46）=8.0, p=.007,　偏η^2=.149）。

トピックをスイッチするということは、その行為により、ストーリーの中での流れが変わる可能性を秘めているという意味で、聞き手にも話し手自身にも注意が喚起させられる。H/Nのように目の前に描写する絵がある場合、聞き手も話し手もある程度情報を共有していることから、トピックをスイッチさせる際に、ゼロ照応を用いたとしても、行動の主体が誰であるか想像がつく。それに比べ、T/Tでは、聞き手は話し手の発する言葉のみを頼りに、ストーリーの内容を理解していかないといけない。今回、聞き手に質問などは控えてもらったこと、また話し手にもできるだけ独話の形で話をしてもらったので、話し手は聞き手になるべく分かりやすく話を伝えていかなければならないという意識も働いたはずである。

前述のとおり、話者がトピックスイッチさせたのは、主要人物について再導入させた場合が多かったのであるが、その中でもゼロ照応を使用した場合はKHの3例を除き、すべて主要人物へのトピックスイッチであった。Clancy（1992）の提唱する‘ellipsis for hero’ストラテジーによると、トピックが主要登場人物にスイッチされる際に、ゼロ照応が使用されるということであるが、今回のデータからもその傾向がうかがえる。

トピックの連続的言及

トピックの連続性の最も高い文脈では、トピックスイッチのコンテクスト同様、「が」、「は」、「も」、助詞の脱落、ゼロ照応の5つの言語形式の使用が見られた。その5つの言語形式の使用割合の平均値を表4に示す。

ゼロ照応

トピックを連続的に言及する際、最も多く使用された言語形式はゼロ照応であった（全グループ平均使用割合H/N：90.1%、T/T：76.1%）。タスクごとに分散分析で5グループを比較した結果、H/Nではグループ間での有意差は見られなかったが、T/Tタスクでは、グループ間でのゼロ照応の使用割合に有意差が見られた（F（4,46）=4.3, p=.005,偏η^2=.271）。シェフェーの多重比較の結果、EMグループのゼロ照応産出割合（61.7%）がNSグループ（p<.05）とKHグループ（p<.05）に比べて有意に低いのが分かった。

H/NとT/T間で、被験者内比較を行ったところ、タスク間でゼロ照応の使用割合に有意な差があることが分かった（F（1,46）=42.7, p=.000, 偏η^2=.481）。

この結果も、トピックスイッチの文脈におけるゼロ照応同様、聞き手と話し手の共有する事象について話をする場合と、話し手のみが持っている情報を用いてストーリーを構築していくことにより差が出たのではないかと思われる。ゼロ照応の使われた指示対象物について、詳しく見て行ったところ、ゼロ照応は主人公について言及する場合が多かったのではあるが、非主要人物についても連続した言及をする場合は、ゼロ照応を使用し、差別化を図っていないのが分かった。

表4 トピック連続的言及に使われた言語形式: H/N、T/Tタスク（数値は使用形式ごとの割合の平均値）

NP+	が	は	も	助詞脱落	ゼロ照応
英語中級					
H/N	4.9	6.3	0	1.7	87.1
T/T	20.7	13.0	1.0	3.6	61.7
韓国語中級					
H/N	3.0	11.4	0	0	85.6
T/T	10.1	8.2	0.6	1.0	80.0
英語上級					
H/N	4.8	7.1	0	0	88.1
T/T	10.1	9.5	0.2	1.6	78.6
韓国語上級					
H/N	1.6	3.6	0	0	94.8
T/T	9.3	8.5	0	1.4	80.7
母語話者					
H/N	0.7	4.1	0	0	95.1
T/T	7.2	8.9	1.1	1.8	80.9
合計平均値					
H/N	3.0	6.5	0	0.4	90.1
T/T	11.7	9.7	0.6	1.9	76.1

注　H/NはHere and Now, T/TはThere and Thenのタスクを表す

フォローアップ面接の結果、タスクごとのゼロ照応の使用・不使用は意識的に行ったものではないことが分かった。共通の意見として得られたことは、映画を思い出しながら話をする場合は、頭の中で考えながら話すために、いかに複雑な内容の話を相手に分かりやすく話すかということに神経を集中させていたということである。それに比べ、絵を見ながら話す場合は、話し手の負担が大幅に軽減され、深く考えることもなく自然に話をすることができたという意見が大半を占めていた。このフォローアップ面接の結果も考慮すると、聞き手と共有していない事象について言及し、ストーリーを発展させていくというタスクを行う場合、話者は聞き手への理解を促進させるために完全名詞句を使用する結果となったと言える。[8]

受身表現

受身表現については、グループごとの使用頻度を報告する。[9]表5に受身表現の出現頻度を示す。

表5に見るように、受身表現の使用頻度は学習者の日本語能力と比例していることが分かる。また、英語話者と韓国語話者で、その使用数に大差がないことから、本研究では、学習者の母語は受身表現の産出数に影響を及ぼしていないと言える。NSと日本語学習者では、異なるパターンが見つかった。NSの発話では、タスク間での受身表現産出数の差はほとんど見られなかったのであるが、学習者では上級話者の場合、母語に関わらずH/Nの方がT/Tより受身表現を多数引き出した。[10] このことは、T/Tのタスクの複雑性が関連していると考える。T/Tの場合、物語の出来事を記憶に頼りながら出来るだけ着実に話そうとするために、主人公を中心に話すという「立場志向」から、ますます離反してしまい、ストーリーの局面ごとに起こった出来事を、事象を中心に話して行ったのではないかと推測できる。

表5　H/N、T/Tタスクにおける受身表現の出現頻度

	H/Nタスク	T/Tタスク	合計
英語中級話者 (N=11)	4	9	13
韓国語中級話者 (N=10)	6	4	10
英語上級話者 (N=10)	18	9	27
韓国語上級話者 (N=10)	18	7	25
日本語母語話者 (N=10)	25	24	49

注　H/NはHere and Now, T/TはThere and Thenのタスクを表す

T/Tよりは学習者による受身形の産出数が多かったものの、H/NにおいてもNSと学習者では差が出たが、これは、物語談話を構成する際の思考の違いから来ているのではないかと考えられる。 先述のとおり、日本語話者はストーリー構築をする際、特定の人物（主人公）に視点をおき、その角度から物語を語るのに対し、英語・韓国語話者の場合は、場面ごとの事象に焦点を置き、出来事を中心に話を進めていく傾向がある。中級話者での受身表現産出が低かったのは、言語能力の限界が考えられるが、上級レベルでの受身表現が母語話者より下回っているのは、やはりこの思考パターンの違いが、物語発話の構成に影響を及ぼした可能性があると言える。実際、NS、

学習者データで見られた受身表現の多くは、主人公が非動作主となっており、トピック性の高い（主語の）位置についたまま話が進められて行っている。NSの場合、主人公に視点をあてて話を進めていくため、主人公を非動作主とした受身表現が頻繁に見られるのに対し、学習者の場合、場面ごとの主体に焦点をあてる傾向があるため、受動文ではなく能動文が産出される傾向が多いのが分かった。このことは、英語話者でも、韓国語話者でも確認できた。

また、受身表現についてフォローアップ面接で質問したところ、学習者は日本語能力レベルに関わらず、受身表現の機能（迷惑受身）自体は理解できていたものの、物語談話の中で、受身表現が動作の受動者である主要登場人物をトピックとして維持するために使用されるということを認識していないことが分かった。

以上の結果から、話者の思考や物事の捉え方が大きく影響を及ぼす受身表現などにおいては、母語での傾向が（ここでは）負の転移として現れたと言える。Slobin（1991）が言うように、学習者が長年培った「思考パターン・コンセプトなど」は談話構成に大きな影響を及ぼし（‘thinking for speaking’）、このような「概念的転移」は学習者の第二言語の能力が上がったとしても、中々排除することはできない難しい問題の一つであると思われる。Odlin（2005）が論じているように、言語の相対性（言語固有の構造が認知に及ぼす影響）と転移の関係は極めて重要なものであり、「概念の転移」を裏付ける証拠が増えつつあるが（Jarvis 1998; Odlin, 2005参照のこと）、未だ詳しく解明されていない分野であり、今後の更なる研究が望まれる領域である（Odlin私信、2005年3月28日）。

まとめ

本研究では、話者がストーリーテリングにおけるトピック管理（トピック導入・維持）をしていく過程で、学習者の母語、タスクの複雑性及び学習者の日本語能力が及ぼす影響を見てきた。指示対象を導入、継続させていく際に使われる言語形式と受身表現の使用について調べたのであるが、研究設問1で挙げられた、トピックの管理能力の発達パターンについては、学習者の母語により影響を受ける可能性があることが再確認された。具体的に言うと、日本語と類型論的に類似した韓国語を母語に持つ話者は、中級レベルで既に、トピック導入の際の「が」の適切な使用や、トピック連続性の高いコンテクストでのゼロ照応の使用ができているのに対し、英語話者の場合、中級ではコンテクストに関わらず「は」の過剰般化が見られたが、日本語能力レベルの向上とともに、母語話者のトピック管理パターンに近似してくることが分かった。しかし、受身表現の使用については、動詞の活用を伴うということもあってか、英語話者だけではなく、韓国語話者も中級レベルでは、産出数が少なく、上級レベルに至っても、学習者の母語に関係なく、ＮＳと比べると受身表現の合計産出数が半数に満たなかった。これは、日本語では、主要登場人物に焦点を当て、その人物を中心にストーリーを構築していくため、受身表現を使用することで主人公を主語の位置で保持するのに対し、英語・韓国語では起こった事象を中心に談話を展開していくため、主人公ではなく、その時々に起こった事象の主体をトピックの位置に持ってくるという違いに起因しているのではないかと思われる。すなわち、学習者の母語での物語を構築していく際の概念がＬ２での談話構成に影響を及ぼしたと考えられる。

二つ目の研究設問にあった、タスクタイプの違いによるトピック管理パターンへの影響であるが、助詞の使用に関しては、韓国語話者においては、タスク間での差は見

られなかったが、英語話者においては、中級レベルでタスクによる違いが見られた。英語話者においては、T/Tのタスクでのナラティブの方が、助詞の脱落が減少した。一方、トピック性の高い文脈におけるゼロ照応の使用に関しては、学習者・日本語母語話者グループ全てにおいて、H/Nタスクでの方がT/Tタスクよりも使用割合が高かった。これは、H/Nでのタスクが、目の前にある話を描写していくということで、トピックを省略しても聞き手に不理解をもたらすという危険性が低いためだと思われる。また、T/Tタスクにおけるゼロ照応の使用割合は、特に英語話者中級レベルで顕著に低いことが分かった。

トピックの連続性に関連する受身表現の出現数のタスク間での差異に関しては、母語話者と学習者で異なるパターンが見受けられた。母語話者による受身表現の出現数がタスクタイプに影響を受けないのに対し、学習者では、上級レベルにおいても、H/Nタスクの方がT/Tタスクよりも受身表現をより多く抽出することとなった。これは、T/Tの場合、H/Nとは違い、記憶に頼りながら、起こった出来事を確実に伝えようとするため、特定の人物を中心に話す「立場志向」から離れ、起こった出来事、その時々の登場人物を中心に話を展開していったのではないかと考えられる。

以上のとおり、母語の違いやタスクの複雑性が、Ｌ２としての日本語でのストーリー構築におけるトピック管理に影響を及ぼすこと、また母語話者と学習者では、タスクの複雑性がナラティブ談話に及ぼす影響が異なるということも明らかになった。

今後の課題

本研究において、結果の解釈にあたり限界があった。本稿では学習者グループ比較により、母語からの転移を示唆したのであるが、Jarvis (2000) が指摘するように、学習者言語、母語話者によるベースラインデータに加え、学習者の母語データも考慮するのが理想的である。

また、データ抽出に使用した両タスクで、授受表現の産出が少数しか見られなかったため、分析対象としなかった。登場人物の多様性、物語の流れなどを考慮し、この２つのストーリーをデータ抽出のため使用したのであるが、今後の課題として、トピックの連続性・視点に大きな関連をもつ授受表現も抽出できるようなストーリーを選出する必要性があると思われる。

母語話者全員にはフォローアップ面接を行う事ができたが、学習者からは1グループ2～3名ずつしかインタビューデータが集められなかった。学習者に意識調査をすることにより、学習者の習得の現状を捉える一助となるだけではなく、指導の指標も定めるのにも役立つことが見込まれるので、今後の課題としたい。

注

1 「私は今朝リンゴを食べた。∅ おいしかった。」のように、既出の名詞（この場合「リンゴ」）を照応する際、照応詞を使わずに言及することをゼロ照応と言う。

2 SOPIはOPIと違い、対話者が存在せず、学習者がテープに向かって時間内にタスクを行い、判定者がテープを聞き、学習者の外国語運用能力レベルを事後判定するというものである。今回も録音する者がそばにいたものの、相づちなどを返さず、学習者にテープを相手とみなして、話すように指示を与えた。

3 映画は12分であったため、記憶力の限界のことも考え、ちょうど話の切れ目であった最初のほぼ6分までに編集し、そのストーリーを話してもらった。

4　例えば、話者Aの「指示対象導入」の回数が10回で、そのうち「ＮＰ＋が」の出現頻度が5回、「ＮＰ＋に」の出現頻度が2回であったとすると、この話者の「ＮＰ＋が」の使用割合は50%、「ＮＰ＋に」の使用割合は、20%となる。

5　この主張に関して詳しくは、Kellerman (1995) を参照のこと。

6　会話データなどでは、無助詞の使用が起こることが報告されているが (長谷川 1993、丸山　1996) 、ナラティブという独話のコンテクストでは、新情報には通常「が」の使用が規範とされている (Hinds,　1984) 。本研究では、聞き手からのフィードバックを最小限にし、モノローグ調で話してもらったためか、NS、韓国語話者のH/Nデータでは、助詞脱落は見られなかった。

7　両タスクでのカイ二乗検定の結果は以下のとおりである。
H/Nタスク
英語中級：$\chi 2$ (1) =33.6 , $p < 0.001$、韓国語中級：$\chi 2$ (1) =67.9, $p < 0.001$ 英語上級：$\chi 2$ (1)=133, $p < 0.001$、 韓国語上級：$\chi 2$ (1) = 64.5, $p < 0.001$、母語話者：$\chi 2$ (1) = 10.6, p =0.001
T/Tタスク
英語中級：$\chi 2$ (1) =97.0 , $p < 0.001$ 、韓国語中級：$\chi 2$ (1) =62.6, $p < 0.001$ 、英語上級：$\chi 2$ (1) =115, $p < 0.001$ 、韓国語上級：$\chi 2$ (1)= 92.6, $p < 0.001$、母語話者：$\chi 2$ (1) = 59.8, $p < 0.001$

8　この結果は、幼児の第一言語としての韓国語習得を調査したClancy (1997) の研究結果とも重なる。

9　授受表現の産出も調べたが、NSのデータでも産出数は極めて少なかったため、今回は分析対象から外した。

10　EMに関しては、T/Tの方が、受身形の産出数が多くなっているが、それは1名の被験者がT/Tで例外的に受身形を6回産出したことによる。

謝辞

本研究は、学術振興会科学研究費 (課題番号15520330) の助成を受けたもので、本稿の執筆段階では、東京外国語大学大学院GP (「国際基準に基づく先端的言語教育者養成」) の補助も受けた。貴重なコメントをくださったJALT編集者及び査読者各位に感謝の意を表すとともに、統計処理について詳しくご指導くださったRusan Chen氏にもお礼を申し上げたい。

引用文献

Andersen, R. (1977). The impoverished state of cross-sectional morpheme acquisition/accuracy methodology. *Proceedings of the Los Angeles Second Language Research Forum* (pp. 308-319). Los Angeles: Department of English, University of California at Los Angeles.

Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition, 12*(1), 43-64.

Clancy, P. (1985). The acquisition of Japanese. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1* (pp. 373-524). Hillsdale, NJ: Lawrence Erlbaum Association.

Clancy, P. (1992). Referential strategies in narratives of Japanese children. *Discourse Processes, 15,* 441-467.

Clancy, P. (1997). Discourse motivations for referential choice in Korean acquisition. In H. Sohn & J. Haig (Eds.), *Japanese/Korean linguistics: Vol. 6* (pp. 639-659). CSLI Publications.

土井利幸・吉岡薫.(1990).「助詞の習得における言語運用上の制約―ピーネマン・ジョンストンモデルの日本語習得研究への応用―」、*Proceedings of the conference on second language acquisition and teaching* (pp. 23-33). International University of Japan.

Fuller, J.W., & Gundel, J.K. (1987). Topic prominence in interlanguage. *Language Learning, 37*, 1-18.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed), *Topic continuity in discourse: Quantitative cross-language studies*. Amsterdam: John Benjamins.

Gundel, J.K. & Tarone, E. (1983). "Language transfer" and the acquisition of pronominal anaphora. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 281-296). Rowley, MA: Newbury.

Harwell, M.R., Rubinstein, E.N., Hayes, W.S. & Olds. C.C (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics*, 17, 315-339.

長谷川ユリ. (1993).「話し言葉における「無助詞」の機能」『日本語教育』、80号、158-168.

Hinds, J. (1983). Topic continuity in Japanese. In T. Givón (Ed.), *Topic continuity in discourse: Quantitative cross-language studies* (pp. 43-93). Amsterdam: John Benjamins.

Hinds, J. (1984). Topic maintenance in Japanese narratives and Japanese conversational interaction. *Discourse Processes, 7*, 465-482.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning, 35*(2), 141-163.

Ishikawa, T. (2005). Investigating the relationship between structural complexity indices of EFL writing and language proficiency: A task-based approach. *JACET Bulletin, 41*, 51-60.

Jarvis, S. (1998). *Conceptual transfer in the interlingual lexicon*. Bloomington, IN: IULC Publications.

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning, 50*, 245-309.

Jin, H. G. (1994). Topic-prominence and subject-prominence in L2 acquisition: Evidence of English to Chinese typological transfer. *Language Learning, 44*(1), 101-122.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere. *Annual Review of Applied Linguistics*, *15*, 125-150.

Kim, H. (1989). Nominal reference in discourse: Introducing and tracking referents in Korean spoken narratives. *Harvard studies in Korean linguistics, 3,* 431-444.

金慶珠.(2001).「談話構成における母語話者と学習者の視点　―日韓両言語における主語と動詞の用い方を中心に―」、『日本語教育』、109号　60-79.

Li, C. & Thompson, S. (1976). Subject and topic: A new typology of language. In C.N., Li (Ed.), *Subject and topic: Symposium on subject and topic, University of California, Santa Barbara* (pp. 457-489). New York: Academic Press.

丸山直子.（1996）.「助詞の脱落現象」、『言語』、25-1:74-80

Master, P. (1997). The English article system. Acquisition, function, and pedagogy. *System*, *25*, 215-232.

Mayer, M. (1969). Frog, where are you? New York: Dial Press.

水谷信子.（1985）.『日英比較話しことばの文法』　東京:くろしお出版.

Nakahama, Y. (2003a). Development of referential management in L2 Japanese: A film retelling task. *Studies in languages and cultures, 25*(1), 127-146. Graduate School of Languages and Cultures, Nagoya University.

Nakahama, Y. (2003b). Cross-linguistic influence on the development of referential topic management in L2 Japanese oral narratives. Unpublished Ph.D dissertation, Georgetown University, Washington, DC.

Nakamura, K. (1993). Referential structure in Japanese children's narratives: The acquisition of wa and ga. In S. Choi (Ed), *Japanese/Korean linguistics*: Vol. 3 (pp. 84-99). Stanford: CSLI Publications.

Odlin, T. (2002). Language transfer and cross-linguistic studies: Relativism, universalism, and the native language. In R. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 253-261). New York: Oxford University Press.

Odlin, T. (2005). Cross-linguistic influence and conceptual transfer: What are the concepts? *Annual Review of Applied Linguistics*, *25*, 3-25.

Polio, C. (1995). Acquiring nothing? The use of zero pronouns by nonnative speakers of Chinese and the implications for the acquisition of nominal reference. *Studies in Second Language Acquisition*, *17*, 353-377.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, *45*(1), 99-140.

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27-57.

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). Blackwell Publishing.

Robinson, P. (2004). Comprehension, cognitive complexity, and task-based language production and acquisition. In D. Smith, S. Nobe, P. Robinson, G. Strong, M. Tani, & H. Yoshiba (Eds.), *Language and comprehension: Perspectives from linguistics and language education* (pp. 187-240). Tokyo: Kuroshio Publishing.

Rutherford, W. (1983). Language typology and language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 358-370). Rowley, MA: Newbury House.

Sakamoto, T. (1993). On acquisition order: Japanese particles WA and GA. *Proceedings of the 3rd Conference on Second Language Acquisition and Teaching* (pp. 105-122).International University of Japan.

Slobin, D. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics*, 1, 7-25.

田中真理. (1996).「視点・ヴォイスの習得－文生成テストにおける横断的および縦断的研究－」、『日本語教育』88号 104-116.

田中真理. (1997).「視点・ヴォイス、複文の習得要因」『日本語教育』92号107-118.

田代ひとみ. (1995).「中・上級日本語学習者の文章表現の問題点―不自然さ・分かりにくさの原因を探る―」、『日本語教育』85号 25-37.

Tomlin, R. (1990). Functionalism in second language acquisition. *Studies in Second Language Acquisition*, *12*, 155-177.

遠山千佳. (2005).「助詞「は」に関する第二言語習得研究の流れと展望－探索的研究と演繹的研究の枠組みから－」、『言語文化と日本語教育』、日本言語文化学研究会102-121.

渡辺亜子. (1996).『中・上級日本語学習者の談話展開』、東京:くろしお出版.

Yanagimachi, T. (2000). JFL learners' referential-form choice in first/through third-person narratives. *Japanese Language Education around the Globe, 10,* 109-128.

Reviews

Talk in Two Languages. Joseph Gafaranga. Hampshire: Palgrave Macmillan, 2007. xi + 225 pp.

Reviewed by
Tim Greer
Kobe University

By definition, bilingual people can and do talk in two languages. Of course this means that sometimes they need to switch from one language to the other, and for many people this can happen repeatedly within a single conversation. This phenomenon has been well documented, leading to a variety of terms and definitions including code switching, language alternation, borrowing, code mixing, and simply bilingual interaction. In this book Joseph Gafaranga does not attempt to confuse the issue by adding any new theories of language alternation to this already extensive list; instead he pulls together the previously published findings and provides a suitably critical overview of the field. He bases his discussion on a series of case studies taken from his own corpus of Kinyarwanda-French code switching collected from talk recorded among Rwandan refugees in Belgium. His study applies what we know about "talk in two languages" and sketches out the challenges it presents to bilingual people in real world situations.

The book is targeted firmly at an academic audience, particularly those who are conducting research in bilingualism. As such, it is not a volume that all language teachers will be interested in; it will be of most benefit to those already familiar with some of the approaches to investigating code switching.

The strength of Gafaranga's book is in its interdisciplinary coverage of the topic. He identifies the two key research approaches as grammatical and interactional, and links them through the recurring theme of searching for order. Taking his cue from ethnomethodology, he defines order as "the very possibility of social action" (p. 3), contending that the fundamental question behind any study of code switching is how speakers manage and make sense of bilingual interaction, despite its apparent disorderliness. While this is a well-founded basis for the discussion, Gafaranga's glossing of order as

"possibility" at times proves to be a somewhat limited depiction of a concept that is crucial to his argument. Order is produced, situated, and occasioned by speakers themselves, and can be found across groups in recurrent interactional resources (Psathas, 1995). To this end, readers should keep in mind that the search for social order in code switching is ultimately an attempt to account for bilingual practices as a systematic collection of interactional resources.

Gafaranga begins by examining "Quasi-theories of language alternation" (Chapter 2), documenting the ways lay people refer to language alternation in nontechnical terms and noting that bilinguals often belittle their own code switching. Importantly, he also outlines some of the shorthand terms that linguists have introduced into the literature, which have been responsible for much of the confusion about what code switching is.

He then looks in further detail at the two main strands of language alternation research: grammatical approaches (Chapters 3 and 4) and interactional approaches (Chapters 5 to 7). He begins by demonstrating from a grammatical perspective how language alternation can be seen as orderly, outlining two general patterns—insertional code switching and alternational code switching. The former adopts an asymmetric view of the two languages, assuming that one is embedded into the other, whereas the latter maintains that the two languages are relatively equivalent. By way of example, consider the following utterance from my own study of Japanese-English code switching (Greer, 2007):

> "That will be one hundred yen *desu kedo*."

Such a sentence can be viewed as either basically English with a little bit of Japanese inserted at the end, or as two distinct languages alternating at the point where they overlap. Gafaranga goes on in Chapter 4 to apply these grammatical models to a case study taken from Kinyarwanda-French language alternation.

Next he moves away from purely linguistic explanations of code switching in order to examine code switching from a socio-functional viewpoint. Chapter 5 offers a comprehensive coverage of socially motivated, identity-related explanations including diglossia, the interactional sociolinguistic notion of "we/they" codes, and Myers-Scotton's Markedness model. Chapter 6 then discusses the organizational perspective, as put forward by Conversation Analysis, the approach which sits best with Gafaranga's pursuit of orderliness. Here he reviews the work of Auer as well as his own re-specification of the notion "language" as "medium."

Chapter 7 applies these socio-functional approaches to another aspect of the author's corpus, the use of language alternation for reporting direct speech. Rather than representing an accurate rendering of exactly what someone said, he maintains that language choice can act as either a depictive or supportive element in a bilingual storytelling sequence, serving to juxtapose a reported speaker's speech against the teller's own utterance, regardless of the language in which the talk was originally conveyed. In other words, when a bilingual speaker is telling someone what a third person said, the reported speech may not necessarily be delivered in the language in which it was originally spoken. A bilingual person might choose to use Japanese to tell a story and then switch to English to quote one of the characters, even if that character originally spoke in Japanese. The function of code switching in this case is to create a distinction between the speaker as narrator and the speaker as character.

Finally, in Chapter 8 Gafaranga uses all these findings to address a real-life issue—language shift among the Rwandan community in Belgium. This case study distinguishes patterns of language choice used by this refugee community which are potentially contributing to the loss of their children's minority language. The importance of this chapter is in demonstrating how these "pure" linguistic theories and socio-pragmatic approaches can be used to address real-life issues, such as language loss within a given community.

This link between theory and application is the most worthwhile aspect of Gafaranga's research. While the book gives a clear and concise overview of the most significant research in the field, the analysis does not simply stop there. By demonstrating how these theories can be applied to his own situation, Gafaranga provides a series of inspiring exemplars which other researchers could use to frame their own investigations on bilingual interaction. Although it is sometimes difficult to follow the Kinyarwanda data, there is still much that applied linguistics researchers can take away, especially those interested in bilingualism. However, much of Gafaranga's discussion focuses on interactants who have been bilingual since early childhood, rather than late bilinguals. Teachers who are interested in the way that Japanese students of English mix their languages may discover that some aspects of Gafararanga's discussion are difficult to apply to classroom contexts; ultimately the focus of this book is not on language learning but on bilingual language use. The study's original contribution lies in its attempt to put forward an interdisciplinary perspective on code switching, offering interested researchers an in-depth overview of the subject, as well as some thought-provoking directions for future investigation.

References

Greer, T. (2007). *Accomplishing identity in bilingual interaction: Codeswitching practices among a group of multiethnic Japanese teenagers.* Unpublished doctoral thesis. University of Southern Queensland, Australia. Available online at http://adt.caul.edu.au

Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. Thousand Oaks: Sage.

Idioms: Description, Comprehension, Acquisition, and Pedagogy. Dilin Liu. New York: Routledge, 2008. xv + 208 pp.

Reviewed by
Ian MacLean
Kansai Gaidai University

When I was told recently that a relative had finally *kicked the bucket*, after a long struggle, I felt a mixture of grief and relief. *Kicked the bucket* may not always be an appropriate way to convey this information. *Passed away* or *is no longer with us* are gentler and more sensitive, but in this case, *kicked the bucket* carried a descriptive force and a metaphorical vividness that the other two expressions lacked. *Kicked the bucket* conveyed the raw reality of the act of dying to me. Idioms are fundamental units of meaning that add color and expression to the language of native speakers; however, they can be difficult for L2 learners, and teachers are often unsure how to approach the inclusion of idioms in a language program. Dilin Liu's book addresses academic and instructional concerns regarding idioms and can serve capably as a textbook in an educational linguistics or a methodology course at either an undergraduate or graduate level. However, although the book is filled with carefully considered pedagogical suggestions, it is not a "how to" ideas book on teaching idioms in the classroom.

Idioms: Description, Comprehension, Acquisition, and Pedagogy is divided into three broad sections. Part I: "Idioms and their Use" offers a comprehensive review of existing research on idioms and how various scholars define them. Part II: "Idiom Comprehension" addresses the research on idiom processing and comprehension, while in Part III: "Acquisition and Pedagogy," Liu reviews the literature on idiom acquisition and its impor-

tance in language programs. Each chapter includes questions for study and discussion. In addition, the book has a useful glossary, an annotated list of selected idiom reference books and textbooks, a list of online resources, and an index.

The book is well organized and academically thorough. Liu strives to be comprehensive in considering what idioms are, how they are processed, and what factors affect their acquisition. He also evaluates the importance of idioms in language programs. As colorful, frequently used expressions, idioms are important, but teachers are often unsure which idioms to introduce and how to teach them. Students, on the other hand, fear the incorrect use of idioms and may avoid them completely.

Liu locates his discussion within a conceptual framework of existing research and pedagogical considerations. The explanation of how idioms should be defined and classified is very good in this respect, as well as his consideration of the historical origins of idioms and whether various idioms have counterparts in other languages. He is also careful to point out developmental factors related to acquisition. Research indicates that students are not able to understand metaphorical idioms until a certain age (p. 94). In building his framework, Liu focuses his discussion on how the research can be exploited educationally with L2 learners. He addresses this directly by devoting Chapters 8 and 9 to macro- and micro-pedagogical strategies, pedagogical checklists, for teachers to evaluate the teaching of idioms in a particular program. These strategies provide a bridge between Liu's academic, research-based framework and actual in-class practice.

However, Liu's book falls short in some areas. He belabors obvious points: we are all aware of the fact that idiom processing and comprehension is much more difficult for L2 learners (p. 74), and that research indicates that "familiar idioms are processed much faster and more accurately than unfamiliar ones" (p. 78). In a few areas, he could have provided more useful information as well. For example, his comments on using a corpus in a practical way to select and teach idioms could have been explained in a simple way as Hall and Lee (2006) have done in their article on using search engines for idiom usage research. In Chapter 9, Liu suggests the value of learning the origins of idioms as a way of enhancing student understanding, but offers no practical advice about how to do this, although his annotated list of resources does have potentially helpful references.

Liu's book will appeal to those with a serious interest in idioms. His review of the literature is thorough and carefully considered. Anyone doing research on idioms will find this a very useful place to become oriented and see the lay of the land. It is not a book that will appeal as a classroom-ready

ideas book, though. It isn't light reading and it does drag a bit in places, but for educators wanting a conceptual understanding of how idioms might fit into their language programs, this book could be exactly the shot in the arm they need.

Reference

Hall, C., & Lee, D. (2006). Creating "quick-and-dirty" corpora with search engines. *Essential Teacher, 3*(2), 38-41.

English Language Learning Materials: A Critical Review. Brian Tomlinson (Ed.). London: Continuum, 2008. x + 333 pp.

Reviewed by
Scott Gardner
Okayama University

With this edited collection of essays, Brian Tomlinson continues his investigation of language learning materials, previously recorded in *Materials Development in Language Teaching* (1998, Cambridge) and *Developing Materials for Language Teaching* (2003, Continuum). While those works served primarily as guides for teachers and curriculum planners to create, evaluate, and adapt materials, the present volume examines how materials developers have done so far in answering the call for more student-oriented materials, specifically in English language education.

The first half of the book studies English language learning materials according to type and target: young learners, science and technology, multimedia, and extensive reading (ER) are just some of the variations explored. Many of these studies reach the same general conclusion, namely that other needs besides those of students often dictate what and how classroom materials are used, sometimes to the detriment of learners' progress: "Over-indulgence in multimedia can provide the wrong signals to people in education who believe that multimedia can drive pedagogically sound methodology" (Mukundan, p. 109); "Reading in the sense of ER is not amenable to the kinds of control so beloved by institutions" (Maley, p. 136); "There is a need for a more balanced approach where students' learning needs (*how* students learn) are given equal weighting to their language needs (*what* needs to be taught)" (Mol & Tin, p. 89).

The second half of the book looks at materials use from a geographical perspective, shifting from native English or ESL regions like the UK and US to EFL regions of Europe, Africa, Southeast Asia, and the Middle East. In some of the contexts covered in this section there has been more success with materials than in others, depending on the particular region's needs for English and its cultural or ideological distance from the western, English-speaking world. As an example, Prodromou and Mishan, in their chapter on Western Europe, discuss the problem of "methodological correctness," where methodologies and materials created in native English speaking countries can be tyrannous to nonnative English speaking teachers around the world, who struggle with methodological mandates and cultural generalizations from "above" as they try to teach appropriately for their own environments.

Conversely, other chapters (e.g., Lumala and Trabelsi on Africa) remind us that over-reliance on local, economical, bottom-up materials development may lack "formative or summative evaluation" (p. 227) in that such materials might not be prepared under the scrutiny of trained linguists or teachers. These materials run the risk of being pedagogically ineffective or outdated. Some governments such as Japan and Kenya find their solution in nationalizing their materials and curricula, which improves cultural and pedagogical oversight, but unfortunately limits individual classroom freedom.

The articles in this volume discuss their subjects in a variety of ways, ranging from the anecdotal (Skeldon on EST; Bolitho on Eastern Europe) to the highly analytic and highly quantitative (Cooker on self-access materials; Smiley and Masui on Japan). At times it becomes difficult for the reader to "switch gears" when reading from one chapter to the next. It also confounds attempts by readers/researchers to make point-by-point comparisons among regions or teaching contexts: hard data for one context in one chapter may be completely unavailable in another.

Tomlinson makes a noble attempt to draw all this disparate research together, but he has to walk a thin line at times. In his concluding chapter, for example, as part of his summarized criticism of language learning materials he lists, next to each other, "They are form-focused and control-centred" and "They are Anglo-centric or Euro-centric . . . in their assumptions about the best ways to learn" (p. 320). Going into detail on these two items can be touchy, since in countries like Greece, Russia, and Japan the criticisms of "Anglo-centric" materials specifically include their methodological over-emphasis on open-ended "communication activities" to the exclusion of form-focused exercises that have been part and parcel of language education in such countries for generations, and indeed may be part of the pedagogical

psyche of these cultures. For curriculum planners, it can amount to being between a rock and a hard place.

For Tomlinson and many of the authors in this book, though, the way out of such straits seems to be through *adaptability*. The best materials are malleable for use in a world of different teaching contexts. And adaptability should not be only the domain of materials developers, but of teachers and learners as well. Tomlinson, citing himself in 2006, surmises that "learning would be far more effective" if materials writers "started to cater more to divergent needs and wants, if teachers more readily and confidently adapted materials . . . and if learners [were] encouraged and helped to make more decisions for themselves" (p. 22). Perhaps the finest example of adaptable materials, described in glowing terms by several authors in the book, is the extended reading library. When properly implemented, the ER library essentially floods learners with authentic and interesting material that they can freely choose from.

English Language Learning Materials: A Critical Review will not easily tell teachers in so many words whether the textbook they're planning to use next semester is worthwhile or not. Nor, for example, will it clearly tell English for Science and Technology curriculum developers in China how their materials needs match or differ from those in Poland or Argentina. In a sense Tomlinson's task is futile: there are innumerable ways in which good or bad materials can be used well or poorly, in favorable or unfavorable conditions. However, one overarching concept that is implied in his conclusion is that *teachers* are the fulcrum on which a text either swings or falls flat. Most of the bold-letter points in Tomlinson's concluding list of "negative characteristics" of ELT materials, such as "underestimating learners" and "overuse of the PPP approach" (p. 319), can be offset by dedicated, innovative teachers with a vested interest in the development of their learners.

This idea of the teacher as the pivotal implementer—and even creator—of materials is what explicitly drove Tomlinson's two previous books on this topic, but that point is not made as strongly here, focusing as it does instead on what publishers, technologies, and governments are doing. In that sense it is not as useful a book for practitioners. But *English Language Learning Materials* may still serve as a motivator for teachers to depend less on what their textbooks tell them to do and to draw more on knowledge of their students, of teaching, and of the world they live in.

Japan's Built-in Lexicon of English-Based Loanwords. Frank E. Daulton. Clevedon: Multilingual Matters, 2008. vii + 185 pp.

Reviewed by
Patrick Foss
Kwansei Gakuin University

When the *abereeji* Japanese *suchuudento* says he likes watching *beesubouru* on *terebi*, his English teacher may want to close her eyes in despair. However, Frank Daulton argues in *Japan's Built-in Lexicon of English-Based Loanwords* that this average student's knowledge of the loanwords for *baseball* and *television* and other terms is actually a positive L1 resource for L2 English acquisition. This is bound to be a controversial assertion in certain quarters, as numerous studies have suggested that cognates in general and Japanese *gairaigo* (Western loanwords) in particular do more harm than good in the language learning process. Nevertheless, the evidence presented in this highly readable, well-researched book is likely to convince many otherwise.

The book is divided into four parts. Part 1, "Japan's Importation of English," gives an entertaining, informative history of the assimilation of loanwords into the Japanese language and explains how the number of English loanwords has increased so dramatically in recent years as to now make up roughly 10% of the Japanese lexicon. In one of the book's most interesting sections, Daulton details the phonological, morphological, grammatical, and semantic ways in which English words are transformed into Japanese.

Part 2, "*Gairaigo* and Language Acquisition," looks at the competing views in the literature on the effects of cognates on L2 English acquisition. Where *gairaigo* is concerned, Daulton summarizes numerous studies demonstrating the positive effects of loanwords on spelling, retention, listening comprehension, and recognition. His points are clear and his evidence well presented. That said, he makes a better case for the benefits of cognates where receptive knowledge is concerned. He cites relatively few studies regarding productive knowledge (it appears few have been done) and seems to overreach somewhat on those he does. For example, Daulton explains how Japanese university students with low English proficiency seem to prefer using borrowed words in written English production. He claims this is "evidence that loanwords are pushing their corresponding borrowed words into production, thus facilitating acquisition" (p. 70). The first part of this claim seems valid. However, to use an example he cites himself, if a

learner produces the borrowed words *flower* and *fire* to express the English word *fireworks*, how exactly is acquisition being facilitated? Without reader intervention (and this assumes an understanding that *flower fire* means *fireworks*), isn't it possible that the production of these borrowed words will have a negative effect? While errors are a natural part of the learning process, it's difficult to see how *gairaigo* is helping the learner here, except by simply encouraging production.

Part 3, "The Built-in Lexicons," is the main section of the book. Much has been made over the past two decades of the value of learners mastering the high frequency words of English. Daulton discovered "that nearly half of the 3000 most frequent word families in English have correspondences with common Japanese loanwords" (p. 81). Not only is this an amazing figure, but the research he presents seems to indicate that the quality of this correspondence is quite high. If one accepts the evidence presented in Part 2 as to the beneficial effects of loanwords, even if only where receptive knowledge is concerned, it's hard to escape the conclusion that these loanwords with connections to high frequency vocabulary words have the potential to be quite useful for Japanese learners of English.

Using cognates does involve difficulties, however. Daulton's final section, "Exploiting Japanese Loanword Cognates," looks at the problems involved in working with *gairaigo* (including orthographic and phonological difficulties) and how it's unclear whether knowledge of individual loanwords can directly enable Japanese learners to extend their knowledge of English words inside word families. Despite these concerns, an epilogue devoted to suggestions on the use of *gairaigo* in the classroom and extensive lists of loanwords in the appendices should spark numerous ideas in the minds of instructors. Some readers will want more detailed suggestions, but that would seem to be beyond the scope of this book.

In the end, one walks away from *Japan's Built-in Lexicon of English-Based Loanwords* with the feeling that simply making learners aware of this "built-in lexicon" would be beneficial. As all learners know, acquiring the necessary vocabulary to be proficient in English is a very difficult task. Shouldn't realizing they already have substantial knowledge of these words serve to make this task easier for them and at the same time boost their confidence? Only the most intractable opponents of *gairaigo* will disagree after reading this book. *Japan's Built-in Lexicon of English-Based Loanwords* is highly recommended as a resource both for individual instructors and high school or university English departments.

Professional Encounters in TESOL: Discourses of Teachers in Teaching. Sue Garton & Keith Richards (Eds.). Houndmills, UK: Palgrave Macmillan, 2008. xxvii + 262 pp.

Reviewed by
Andre A. Parsons
Hokkaido University of Education, Hakodate

What do English language (EL) teachers do and why do they do it? This is the profound question that editors Sue Garton and Keith Richards explore in this volume. In order to do so, they have invited EL practitioners from various backgrounds to examine different aspects of professional talk that teachers encounter at different points in their careers. The result is an accessible and thought-provoking read with some practical information and ample resources for further study for beginning EL teachers, especially for those who are interested in discourse analysis.

The book is organized into stages according to a modified version of Huberman's classic description of the teacher career cycle which the editors have renamed as 1) Starting Out; 2) Becoming Experienced; 3) New Horizons; and 4) Passing on the Knowledge. Garton and Richards explain that they omitted Huberman's final stage, *disengagement*, as it falls outside the scope of the book. However, it might have been useful to include such a chapter to view a teacher's career from start to finish as it could give distinct insights into the development of EL teachers throughout their careers. The editors also assume that all EL teachers start out as teacher-trainees, but some teachers find themselves in such a career with no training, as if they "sort of drifted into EFL" (p. 174), to take the words of Harry from Richards' chapter. Where would such a teacher fit in this model?

Although the topics have been organized according to their relevance to each stage in the career model, readers will be able to relate to and learn from most of the chapters in this book. In "Starting Out," Copland's and Hooton's chapters on feedback are useful to both teacher-trainees and teacher trainers. Seedhouse's chapter on classroom discourse encourages teachers to reflect on how they interact with their students and how they can improve this skill.

In "Becoming Experienced," Garton's article on teacher beliefs and the following reflections by Tarnpichprasert are not only valuable to novice teachers, but to experienced ones as well, while Howard's chapter on classroom appraisal may be useful for teachers having to undergo such observations.

Morris-Adams' chapter on informal talk between NS and NNS and how it can be beneficial to students outside the classroom, although of interest to ESL practitioners, may not be so helpful for those teachers in an EFL setting, where opportunities to practice English outside of class are limited.

In "New Horizons," Quirke's chapter on web-based support for teacher development and Mann's chapter on cooperative development and metaphor are pertinent to all levels of experience as they both focus on the important concept of fostering personal and professional growth through community support. Richards' chapter on the experience of establishing a new school may be interesting for those considering such a move as he offers some important lessons necessary to succeed in such a venture.

Finally, in "Passing on the Knowledge," Kuchah describes his experiences in the transition from being a classroom teacher to becoming one of the youngest school inspectors in Cameroon. In particular, he discusses how he had to overcome the challenges of the power relationships in education that exist between the youth and elders in his society. Wharton's chapter provides advice to those who are contemplating taking that daunting, yet important step of writing for publication, while the final chapter of the book by Edge focuses on semantic patterning, action research, and cooperative development, and their interrelationship. Anyone interested in such topics will find Edge's chapter insightful as this appears to be the first attempt to investigate the underlying similarities of those constructs. However, as the concluding chapter, it wasn't effective because it failed to tie the various strands of the book together.

There is one major contradiction in the book, though, highlighted in Wharton's chapter: "Speaking for oneself is an indicator of power, whereas being spoken for by others suggests relative powerlessness" (p. 229). She is referring to the frequent use of teachers as data, but the lack of teachers as authors. The editors do include reflective chapters by EL teachers on their current stages in teacher development, but these are simply reflections, not original research. The overall impact of the book would have been greater had there been more entries written by EL teachers at each particular stage; however, it seems that most practicing EL teachers do not take that step; an issue that should be dealt with in future research.

As a teacher between the "Starting Out" and the "Becoming Experienced" stages, I am very interested in the stories that EL teachers tell and in what can be learned from them. It is through interacting with one another, both dialogically and monologically, that we learn and grow as teachers. For me, this book was quite informative and practical. I especially appreciated Seedhouse's and Garton's chapters as both encouraged me to reflect on my

own teaching and interactions with my students; Quirk's chapter on supporting teacher development on the web was quite valuable in providing an alternative venue for interacting with and learning from other EL teachers, both beginning and experienced; and Wharton's chapter presented useful information on writing for publication.

The question of what a teacher does and why is a valid one that will continue to receive attention in the future. Overall, this book is useful, albeit in different ways, for all EL teachers, both experienced and inexperienced. While those newer to the field may find the book informative and enlightening, those with considerable experience in TESOL may find it of more use in helping them gain a perspective on how they have developed as teachers. Whether you are in the "Starting Out" stage or the "Passing on the Knowledge" stage, this book should prompt you to reflect upon and share your own development as a professional with others.

Information for Contributors

All submissions must conform to *JALT Journal* Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the refereed research journal of the Japan Association for Language Teaching (*Zenkoku Gogaku Kyoiku Gakkai*), invites practical and theoretical articles and research reports on second and foreign language teaching and learning in Japanese and Asian contexts. Submissions from other international contexts are accepted if applicable to language teaching in Japan. Areas of particular interest are:

1. Curriculum design and teaching methods
2. Classroom-centered research
3. Cross-cultural studies
4. Testing and evaluation
5. Teacher training
6. Language learning and acquisition
7. Overviews of research and practice in related fields

The editors encourage submissions in five categories: (1) full-length articles, (2) short research reports (*Research Forum*), (3) essays on language education or reports of pedagogical techniques which are framed in theory and supported by descriptive or empirical data (*Perspectives*), (4) book and media reviews (*Reviews*), and (5) comments on previously published *JALT Journal* articles (*Point to Point*). Articles should be written for a general audience of language educators; therefore statistical techniques and specialized terms must be clearly explained.

Guidelines

Style

The *JALT Journal* follows the *Publication Manual of the American Psychological Association*, 5th edition (available from APA Order Department, P.O. Box 2710, Hyattsville, MD 20784, USA; by e-mail: <order@apa.org>; from the website: <http://www.apa.org/books.ordering.html>). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. A downloadable copy of the *JALT Journal* style sheet is also available on our website at <http://www.jalt-publications.org/jj/>.

Format

Full-length articles must not be more than 20 pages in length (6,000 words), including references, notes, tables, and figures. *Research Forum* submissions should be not more than 10 pages in length. *Perspectives* submissions should not be more than 15 pages in length. *Point to Point* comments on previously published articles should not be more than 675 words in length, and *Reviews* should generally range from 500 to 750 words. All submissions must be typed and double-spaced on A4 or 8.5"x11" paper with line spacing set at 1.5 lines. The author's name and identifying references should appear only on the cover sheet. Authors are responsible for the accuracy of references and reference citations.

Submission Procedure

Materials should be sent as an e-mail attachment in Rich Text Format (preferred) or by post. Postal submissions must include a clearly labeled floppy disk or CD-ROM and one printed copy. Please submit materials to the appropriate editor indicated below.

Materials to be submitted— all of this can be sent by e-mail in one attached RTF file (preferred) or by post. If by post, be sure to include all of the material on a floppy disk or CD-ROM.

1. Cover sheet with the title and author name(s).
2. One (1) copy of the manuscript, with no reference to the author. Do not use running heads.
3. Contact information sheet, including one author's full address and, where available, a fax number and e-mail address.
4. Abstract (no more than 150 words).
5. Japanese translation of the title and abstract, if possible (no more than 400ji).
6. Biographical sketch(es) (no more than 50 words each).
7. Authors of **accepted** manuscripts must supply camera-ready copies of any diagrams or figures and a copy of the manuscript (RTF or ASCII).

Evaluation Procedures

Submissions will be acknowledged within one month of their receipt. All manuscripts are first reviewed by the editorial board to insure they comply with *JALT Journal* Guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to: (1) compliance with *JALT Journal* Editorial Policy, (2) the significance and originality of the submission, and (3) the use of appropriate research design and methodology. Evaluation is usually completed within three months. The first author of a published article will receive one copy of the *Journal* and 10 off-prints with the option to order more at a set rate (contact JALT Central Office for price details). Review authors receive one copy of the *Journal*.

Restrictions

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned. If accepted, the editors reserve the right to edit all copy for length, style, and clarity without prior notification to authors.

Full-Length Articles, Research Forum, and Point to Point Submissions

Please send submissions in these categories or general inquiries to:

jj-editor@jalt-publications.org

Ian Isemonger, *JALT Journal* Editor

Perspectives

jj-editor@jalt-publications.org

Ian Isemonger, *JALT Journal* Editor

Reviews

The editors invite reviews of books, tests, teaching systems, and other publications in the field of language education. A list of publications that have been sent to JALT for review is published monthly in *The Language Teacher*. Please send submissions, queries, or requests for books, materials, and review guidelines to:

jj-reviews @jalt-publications.org

William Perry, *JALT Journal* Reviews Editor

Japanese-Language Manuscripts

JALT Journal welcomes Japanese-language manuscripts on second/foreign language teaching and learning as well as Japanese-language reviews of publications. Submissions must conform to the Editorial Policy and Guidelines given above. Authors must provide a detailed abstract in English, 500 to 750 words in length, for full-length manuscripts and a 100-word abstract for reviews. Refer to the Japanese-Language Guidelines for details. Please send Japanese-language manuscripts to:

jj-editorj@jalt-publications.org

Yoshinori J. Watanabe, Ph.D; *JALT Journal* Japanese-Language Editor

Address for Inquiries about Subscriptions, Author Reprints, or Advertising

JALT Central Office

Urban Edge Building 5F

1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan

Tel.: 03-3837-1630; Fax: 03-3837-1631

(From overseas: Tel.: 81-3-3837-1630; Fax: 81-3-3837-1631)

E-mail: jco@jalt.org URL: http://www.jalt.org

日本語論文投稿要領

*JALT Journal*では日本語で執筆された論文、研究報告、実践報告、書評等を募集しています。文体:一般的な学術論文のスタイルを用い、章立ての仕方や参考文献のデータの書き方などは、Publication Manual of the American Psychological Association (5th ed.)の定める方式に合わせて下さい。不明の場合は、*JALT Journal* の英語論文を参考 にするか、日本語編集者までお問い合わせ下さい。なお、*JALT Journal*の読者は現場の 教師が主なので、特殊な専門用語や統計的手法は、わかりやすく定義するか説明を加えるなどして下さい。原稿: 長さは、参考文献リストも含め18,000字 (書評の場合は 1,500字) 以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷して下さい。手書きの原稿は受け付けません。図表をいれる場合は、*JALT Journal*のページの サイズに合わせて作成して下さい (縮小コピー可) 。

提出するもの:
以下の原稿を電子メールの添付書類、あるいは郵送でお送りください。

- 執筆者の名前と所属機関名を書いた表紙
- 連絡先(住所、電話番号、ファックス、e-mail アドレス)
- 400字以内の和文要旨
- 英文のタイトルと、500〜750語の英文要旨(書評の場合は100語程度の英文要旨)
- 100字以内の執筆者略歴
- MS-Word で保存した本文(マッキントッシュ使用の場合はrtf ファイルで保存したもの)(執筆者は無記名のこと)
- 審査を経て掲載の認められた草稿は、図表などを全て写植版にしたものにして提出すること

査読: 編集委員会で投稿要領に合っているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿が *JALT Journal*の 目的に合っているか、 言語教育にとって意味があるか、 独創性はあるか、研究計画や方法論は適切か等が判定されます。査読は通常二か月以内に終了しますが、特に投稿の多い場合などは審査にそれ以上の時間がかかることがあります。

注意: *JALT Journal*に投稿する原稿は、すでに出版されているものや他の学術雑誌 に投稿中のものは避けて下さい。*JALT Journal*は、そこに掲載されるすべて の論文に関して国際著作権協定による世界初出版権を持ちます。なお、お送りいただいた原稿は返却しませんので、控を保存して下さい。

投稿原稿送り先またはお問い合わせ:

〒102-8554 上智大学 外国語学部 言語学副専攻
JALT Journal 日本語編集者 渡部良典
電話・Fax: 03-3238-4243
jj-editorj@jalt-publications.org

JALT Journal 第 31 巻 第 1 号

2009年4月20日 印刷
2009年5月 1日 発行
編集人 イアン・アイズマンガー
発行人 スティーブ・ブラウン
発行所 全国語学教育学会事務局
〒110-0016 東京都台東区台東1-37-9 アーバンエッジビル5F
TEL(03)3837-1630; FAX(03)3837-1631
印刷所 コーシンシャ株式会社
〒530-0043 大阪市北区天満1-18-4天満ファーストビル301 TEL(06)6351-8795