LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

"I Prefer Not Text": Developing Japanese Learners' Communicative Competence with Authentic Materials

Alex Gilmore

The University of Tokyo, Japan

This quasi-experimental study reports on a 10-month classroom-based longitudinal investigation, exploring the potential of authentic materials to develop Japanese learners' communicative competence in English. Sixty-two second-year university students were assigned to either a control group receiving textbook input or an experimental group receiving authentic input, and their pretreatment and posttreatment levels of overall communicative competence were assessed. Communicative competence was operationalized with a batch of eight different tests: a listening test, a pronunciation test, a C-test, a grammar test, a vocabulary test, a discourse completion task, an oral interview, and a student-student role-play. The results indicated that the experimental group outperformed the control group in five of the eight measures, suggesting that the authentic materials and their associated tasks were more effective in developing a broader range of communicative competencies in learners than the textbook materials. I discuss the pedagogical implications of these findings for language teachers and their learners.

Keywords authenticity, communicative competence, noticing, materials design, classroom-based research, Japanese EFL learners

When learners are given the choice of studying a second language (L2) from either textbooks or authentic materials (such as films, songs, novels, or Web-based sources), they very often "prefer not text," as one of my own students told me in no uncertain terms. The motivating nature of authentic materials has often been noted in the research literature (Bacon & Finnemann, 1990;

This article reports on research conducted for my Ph.D. thesis, submitted to the School of English Studies, University of Nottingham, England. The author would like to thank Zoltán Dörnyei, Adela Gánem Gutierréz, Michael Handford, Tatsuya Taguchi, Michiyo Yamazaki, and three anonymous reviewers for their invaluable comments on earlier drafts of the article.

Correspondence concerning this article should be addressed to Alex Gilmore, Department of English Language, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan. Internet: alexgilmore@mac.com

© 2011 Language Learning Research Club, University of Michigan
DOI: 10.1111/j.1467-9922.2011.00634.x

Freeman & Holden, 1986; González, 1990; Keinbaum, Russell, & Welty, 1986; Little, Devitt, & Singleton, 1989; Little & Singleton, 1991; McGarry, 1995; Morrison, 1989; Peacock, 1997; Swaffar, 1985), and this in itself constitutes a powerful justification for their use in the language classroom. However, this article focuses on a different quality inherent in authentic materials, which is their ability to highlight a wide variety of discourse features and, through this, develop a range of communicative competencies in learners.

Communicative Competence

The spate of interest in communicative competence, first sparked by Hymes's (1972) well-known attack on Chomsky's notion of an "ideal speaker-listener" and his trivialization of performance in the 1970s, seems to have waned in recent years among applied linguists. Although the term has become common currency in both the research literature and in language textbooks, "communicative competence" means different things to different people, and it would therefore seem sensible to begin this article with a clear definition of the construct. The model proposed here is one that has emerged gradually over the last 35 years from the work of a number of researchers (principally Canale, 1983; Canale & Swain, 1980; Celce-Murcia, Dörnyei, & Thurrell, 1995; Clyne, 1979; Hymes, 1972; Leech, 1983; Schmidt & Richards, 1980; Thomas, 1983) and appears to be broadly accepted by experts in the field.[1] It consists of five interrelated components:

1. *Linguistic Competence*: This refers to a speaker's lexical, morphological, orthographical, syntactical, and phonological knowledge of the language. In other words, how to build up morphemes into words and words into clauses and sentences and how to spell them in the written form or pronounce them in speech. It deals only with the literal meaning (or locutionary force) of utterances. This is the type of knowledge that has traditionally been the staple diet of English language teaching (ELT) classrooms and it is important to note that it is not rejected in the current model of communicative competence but rather assumes a lesser role, seen as only one aspect of language proficiency.
2. *Pragmalinguistic Competence*: This refers to a speaker's ability to understand or convey communicative intent appropriately in a given context based on a knowledge of phrases typically used by native speakers to express speech acts such as apologies, requests, refusals, and so on. This kind

of competence therefore describes a speaker's ability to interpret the illocutionary force, or conversational implicature (Grice, 1975), of utterances.

3. *Sociopragmatic Competence*: This refers to a speaker's knowledge of what is socially or culturally appropriate in a particular speech community. This might include an appreciation of politeness and social conventions, taboo topics, and nonverbal factors such as kinesics and proxemics.
4. *Strategic Competence*: This refers to a speaker's ability to exploit verbal or nonverbal communication strategies when communication problems arise, compensating for deficiencies in other competencies. These include four common types:
   - *Avoidance or reduction strategies* such as topic avoidance or message abandonment to try to keep conversation inside areas where the speaker feels in control;
   - *compensatory strategies* such as circumlocution or mime when a word is not known;
   - *stalling strategies* such as using hesitation devices or repetition to hold the turn in conversation while a message is formulated;
   - *interactional strategies* such as asking for repetition or clarification where the speaker makes use of the linguistic resources of other interlocutors to maintain conversation.
5. *Discourse Competence*: This refers to a speaker's ability to produce unified, cohesive, and coherent spoken or written discourse of different genres (Halliday & Hasan, 1989). In writing, this might include the knowledge of the correct layout for a letter or how to use anaphoric reference in a text. In speaking, it would include how to develop a conversation naturally through "topic shading," in which a subtopic from preceding talk is taken up and expanded into the main topic (Bublitz, 1988; Crow, 1983). It could also include knowledge of different generic structures such as narratives, gossip, or jokes (Eggins & Slade, 1997), or discourse intonation (Brazil, Coulthard, & Johns, 1980).

Having settled on the above-described model with its five interrelated components, it may seem as if the major questions associated with this area of research have all but been resolved. However, we still have little understanding of a wide range of issues, including the following:

1. the extent to which the different components of the model can be viewed as separate entities;
2. exactly how the different components interact together, and whether they can be developed independently of each other;

3. the extent to which a focus on the different components of the model in the language classroom can benefit learners' developing communicative competence;
4. the relative importance of each type of competence in successful communication in specific contexts;
5. the relative importance of each type of competence at different stages of language proficiency;
6. how to effectively measure the different components of the model.

Byram (1997) saw the components of the communicative competence model as closely interrelated and largely inseparable from each other, and intuitively this makes sense. Obviously, learners need a certain level of linguistic competence in order to produce speech acts (pragmatic competence) or build up longer stretches of discourse (discourse competence). Similarly, strategic competence can improve the quality of scaffolding (Bruner, 1983) learners receive in their L2 interactions (by indicating to an interlocutor exactly how their utterances need to be reworked to be comprehensible), thereby enhancing their development of linguistic competence.

However, the relationship between the different components is more complicated than first impressions might suggest. For example, pragmatics researchers have observed a "U-shaped curve" in learners' developing pragmatic competence, as they acquire sufficient lexicogrammatical knowledge for negative pragmatic transfer from the first language (L1) to occur (e.g., Kasper & Rose, 1999) or, metaphorically, as they gain enough rope to hang themselves, as it were. Hatch (1978) and Day (1986) noted that, in L1 development, we learn the "discourse frames" for interaction before we learn the language to slot into them, suggesting that, in some respects, discourse competence develops in advance of linguistic competence. Other empirical evidence does little to clarify the relationships among the various components of the model. Bachman and Palmer (1982) developed a battery of tests to measure (using their terminology) "grammatical competence" (morphology and syntax), "pragmatic competence" (vocabulary, cohesion and organization), and "sociolinguistic competence" (sensitivity to register, naturalness, and cultural references). They found that grammatical and pragmatic competencies were closely associated with each other and that sociolinguistic competence was distinct. However, the components they included within pragmatic competence are more commonly associated with linguistic or discourse competence so their results are difficult to interpret within the framework proposed here. Allen, Bernhardt, Berry, and Demel (1988) found no significant differences among grammatical

competence (morphology and syntax), discourse competence (cohesion and coherence), and sociolinguistic competence (sensitivity to register) in their study using factor analysis of test scores. On the other hand, Schmidt's (1983) 3-year longitudinal study of the development of communicative competence in Wes, a Japanese artist living in Hawaii, found that his discourse and pragmatic competence developed significantly, whereas his grammatical knowledge changed very little, suggesting that these components *are* distinct from each other. More recent investigations have focused principally on linguistic versus pragmalinguistic competence and confirm the distinctness of these two components (e.g., Bardovi-Harlig & Dörnyei, 1998; Kasper, 2001a; 2001b; Salsbury & Bardovi-Harlig, 2001).

With respect to the relative importance of different areas of communicative competence, Celce-Murcia et al. (1995) saw discourse competence as playing a central role in the selection and sequencing of words or structures to form unified spoken and written texts. Bachman (1990, p. 103), on the other hand, expanded the notion of strategic competence and saw it as mediating between the communicative goal in a given situation and the language resources available to a speaker. It assesses how best the speaker's competencies can be exploited to achieve the communicative goal, retrieves the relevant items, and plans the execution of the message. It then assesses how well the goal has been achieved. This seems remarkably similar to the "core role" assigned to discourse competence by Celce-Murcia et al.

For the language teaching profession, faced with a pressing need for immediate answers to these issues in order to guide classroom practice, the lack of a clear consensus, highlighted earlier, is somewhat frustrating. However, as a general point of departure, it would seem reasonable to aim at producing learners with competence in all five areas of the proposed model. Therefore, two important practical questions that emerge from this position, and which will be addressed in this article, are as follows:

1. To what extent do current EFT textbooks encourage the development of a broad range of communicative competencies in learners?
2. Would a greater focus on the different components of the model in the classroom lead to demonstrable changes in learners' overall communicative competence?

The Representation of Communicative Competence in Language Learning Materials

Although the quality of ELT textbooks has undoubtedly improved over the last few decades, there is a large body of research to suggest that they often

continue to present learners with an impoverished or distorted sample of the target language to work with, and fail to meet many of their communicative needs (see Gilmore, 2007a, 2007b, for an extensive review of the literature dealing with these issues). There are two main reasons for this state of affairs. First, descriptions of the English language have traditionally centered on sentence-level, lexicogrammatical features and, as we tend to teach only what we understand, textbook materials have also mirrored this bias. Research from the fields of discourse/conversational analysis, pragmatics, and sociolinguistics has provided us with a clearer picture of the characteristics of natural discourse so that we are now better able to evaluate the descriptions on which we base our teaching (McCarthy, 1991; McCarthy & Carter, 1994). Second, language learning materials have typically been based on contrived discourse, invented by writers to illustrate particular points in itemized structural syllabi. This reliance on the intuitions of authors has often produced misleading models of the target language, as it is well established that although people are very good at noticing unusual patterns in their mother tongue, they are highly unreliable when it comes to an awareness of typical speech patterns (Biber, Conrad, & Reppen 1994; Labov, 1966; Sinclair, 1991; Wolfson, 1989).

The obvious solution to these issues and one that is becoming increasingly common is to base language models, partially or exclusively, on authentic discourse, defined here, in the same way as by Morrow (1977, p. 13), as "a stretch of real language, produced by a real speaker or writer for a real audience and designed to convey a real message of some sort."[2] This definition is deliberately broad and could include texts such as nonnative-speaker discourse or motherese, but it is meant to distinguish between language that has been produced to communicate a genuine message and language contrived by material writers to display particular lexicogrammatical items. For example, an imagined textbook dialogue, composed by a writer, would not be termed authentic here because it is not produced contingently by two real speakers, collaborating in real time to construct a conversation together, with attendance to issues such as face (Brown & Levinson, 1987). Nor is it intended to convey a real message but instead aims to display specific language features in an itemized syllabus to the target audience.

Authentic materials, particularly audiovisual samples (Brown & Yule, 1983), offer a much richer source of input in the classroom and have the potential to raise learners' awareness of a wider range of discourse features and are therefore, hypothetically, more likely to encourage the development of a broader range of communicative competencies in learners. This position is based on cognitive theories of second language acquisition (SLA), which see

"noticing" as playing a crucial role in interlanguage development (Batstone, 1996; Schmidt, 1990; Skehan, 1998; Wigglesworth, 2005). As Schmidt (2001, pp. 3–4) put it, "SLA is largely driven by what learners pay attention to and notice in the target language input and what they understand the significance of noticed input to be." Most discussions on the role of attention in L2 learning have so far focused principally on form features, such as morphology and syntax (Schmidt, 2001), but it is hypothesized here that other characteristics of the input, including prosodic or paralinguistic features, can also be acquired by learners when they are encouraged to notice them through careful selection of materials and principled task design.

Purpose of the Present Study

The present study aimed to explore the effects of authentic versus textbook input on learners' development of linguistic, pragmalinguistic, sociopragmatic, strategic, and discourse competencies. It was hypothesized that the richer input provided by authentic materials, combined with appropriate awareness-raising and practice activities, would allow a wider range of discourse features to be noticed by the learners and lead to enhanced development of their overall communicative competence.

Method

In the design of this investigation, criticisms of earlier classroom-oriented studies were taken into consideration in an effort to maximize the validity and reliability of the results. These include a scarcity of research carried out in genuine ELT classrooms (Nunan, 1991, 1996), a bias toward features that are easily observed or measured (Allwright & Bailey, 1991; Johnson, 1995), a lack of equivalence between comparison groups, and research periods in longitudinal studies too short for measurable changes to take place (Alderson & Beretta, 1992; Ritchie & Bhatia, 1996).

Procedure

The study was conducted over a 10-month period, from April 2004 to January 2005, using four intact classes of second-year students from a university in the Kansai area of Japan. Two of the classes were randomly assigned to the control treatment, receiving only textbook input, whereas the other two were assigned to the experimental treatment, receiving predominantly authentic input. Classes were held twice a week in quiet, well-lit classrooms for all groups, each lesson

lasting 90 minutes (a total of 82.5 hours of input over the course of the investigation), and were all taught by one teacher (the author[3]). Quantitative data on the students were collected precourse and postcourse with a batch of eight different tests.[4] All training and testing took place during scheduled classes, except for the International English Language Testing System (IELTS) oral interviews, which were arranged outside of class time, by appointment.

Participants

A total of 62 second-year English-major students, from four intact classes in the university, took part in this quasi-experimental study. Ages ranged from 19 to 22 ($M = 19.2$), with a female-to-male ratio in the classes of 2:1. Participants reported between 7 and 15 years of previous English language instruction ($M = 7.9$) and represented those learners with the highest English language proficiency in the university, with TOEFL scores ranging from 493 to 567 ($M = 514.3$). Students typically had around 9 hours of formal English language input per week at the university, only 3 hours of which involved the experimental or control treatment (other classes focused on reading or writing skills, business English, current affairs, and preparation for the TOEIC, the Test of English for International Communication).

Input Materials

Training for both control and experimental groups focused principally on developing learners' listening and speaking skills, as these were the areas of priority on the "Communicative English Course" in which participants were enrolled. However, the type of input to which students were exposed (the independent variable) differed significantly.

Control Group

The control group worked methodically through two selected textbooks—*Inside English* (Maggs, Kay, Jones, & Kerr, 2004) and *Face to Face* (Fuller & Fuller, 1999)—using the materials and tasks provided by the authors with occasional supplementation from other teaching resource books where it was felt necessary. These particular texts were chosen because they were subjectively judged to contain predominantly contrived texts, designed for pedagogic exploitation, which helped to create a sharp contrast between the input which the experimental and control groups received.[5]

Experimental Group

The experimental group received predominantly (but not exclusively) authentic materials (as defined by Morrow, 1977) throughout the trial, taken from

films, documentaries, reality shows, TV comedies, Web-based sources, home-produced video of native speakers, songs, novels, and newspaper articles. Materials were selected on the basis of their ability to highlight some aspect of communicative competence, along similar lines to those suggested by Celce-Murcia, Dörnyei, and Thurrell (1997) in their "principled communicative approach," and so, at times, textbook resources were utilized where it was considered expedient. The syllabus is summarized in Table 1.

Testing Instruments

Eight different tests were administered in order to assess students' overall communicative competence. The same measures were used for both precourse and postcourse evaluation, making the assumption that the time between tests (36 weeks) would be sufficient to counter any practice effects.

1. Listening test: The listening test used was an IELTS practice test, taken from *Passport to IELTS* (Hopkins & Nettle, 1995, pp. 130–132). It was composed of four separate dialogues, each centered on an Australian female studying at a British university.
2. Pronunciation test: The receptive pronunciation test used was taken from *Speaking Clearly* (Rogerson & Gilbert, 1990, pp. 2–6). Nine sections were included in the test, covering syllable stress, weak forms, individual sound recognition, rhythm, word recognition and catenation, sentence stress, and intonation.
3. C-test: The C-test was adapted from texts taken from four different levels of the Headway Series, published by Oxford University Press (see Appendix S1 in the online Supporting Information for this article). It is similar to a traditional cloze test except that it involves deletion of the second half of every second word, starting and ending with an intact sentence (Klein-Braley & Raatz, 1984). Dörnyei and Katona (1992) compared the C-test with a standard cloze test and found it to be a superior measure of general language proficiency, particularly with more homogeneous groups (as was the case with the participants in this study).
4. Grammar test: The grammar test used was taken from *English Grammar in Use: Intermediate Level*, 2nd edition (Murphy, 1994, pp. 301–309). It consisted of 15 sections and a total of 121 multiple-choice items. A wide range of grammatical structures are covered in the test: present, past, present perfect, and future tenses; modal auxiliaries; conditionals; passives; reported speech; questions and auxiliary verbs; –ing and infinitive constructions; articles and nouns; pronouns and determiners; relative clauses; adjectives and adverbs; conjunctions; and prepositions.

Table 1 Summary of syllabus for experimental group

Theme	Content
1. Dictionary skills	Using a monolingual dictionary effectively
2. Listening to NSs of English	Stress-timing, linking, weak forms, accents from around the world (Ellis & Sinclair, 1989, pp. 56–57), *Tom's Diner* (Suzanne Vega), *This is the house that Jack built* (children's verse), phonemic charts
3. English pronunciation and intonation	Scenes from *My Fair Lady* (George Cukor), "as + adj. + as + noun" expressions, acting out scenes
4. Circumlocution strategies	What to do when you don't know a word (Ellis & Sinclair, 1989, p. 39), describing unfamiliar objects, miming activities
5. Conversational repair strategies	Taking control of a conversation (Ellis & Sinclair, 1989, p. 63), giving directions and practice using conversation strategies
6. Hesitation devices and British sociopragmatic conventions	Scenes from *Big Brother* (UK reality show), practice using common hesitation devices, introductions in English, colloquial expressions, HW assignment—student Big Brother audition tapes, *in case* vs. *so that* (Naunton, 1994, pp. 32–33)
7. Starting conversations in English (sociopragmatic and pragmalinguistic conventions)	Extracts from *Around the World in Eighty Days* (Verne, 1873) and scenes from *Around the World in 80 Days* (BBC TV series), strategies and expressions for opening up conversations with strangers, role-play activities, HW assignment—start a conversation with a stranger
8. Closing conversations in English (sociopragmatic and pragmalinguistic conventions)	Scenes from *Annie Hall* (Woody Allen), scenes from *Louis Theroux's Weird Weekends* (BBC TV series), practice closing conversations
9. Discourse intonation	Sentence stress, "telling" and "referring" (fall/fall-rise) tones, tone units, tonic stress (Bradford, 1988, pp. 5–17)
10. Developing conversations in English	Interview with a musician (Falla, 1994, pp. 20–24), extract from *Polite Fictions* (Sakamoto & Naotsuka, 1982, pp. 80–87), strategies for developing conversation in English, transition relevance places, topic shift, paralinguistics of turn-taking, practice developing conversations

(*Continued*)

Table 1 Continued

Theme	Content
11. Listener responses (reactive tokens) and ellipsis in spoken English	Scenes from *Secrets and Lies* (Mike Leigh), showing interest, surprise, understanding, agreement, and so on in English, ellipsis, colloquial language, practice using reactive tokens in conversation, discussing adoption, role-play—Hortense meeting her mother for the first time
12. Oral narratives and register in English	Scenes from *Reservoir Dogs* (Quentin Tarantino), conversational story-telling skills (Jones, 2001, pp. 155–163), structure of oral narratives, strategies for making stories interesting (use of historic present, exaggeration, pitch range, body language), formal/informal register, colloquial language, role-play scenes, HW assignment—telling a personal story in an interesting way, taboo words
13. Formal and Informal registers in English (pragmalinguistic conventions)	Scenes from *Fargo* (Joel and Ethan Coen), scenes from *Fawlty Towers* (BBC TV series), article from *The Guardian* on the British class system, "What Class Are You?" quiz (http://www.pbs.org/peoplelikeus/games/index.html), features of formal and informal discourse in English, using intonation to show politeness, role-plays—checking in and making complaints in a hotel (politely or impolitely)
14. Listening to NSs of English II	Short documentary films from *Video Nation* (BBC Web site): http://www.bbc.co.uk/videonation/; transcribing NS English, Video Nation presentations
15. Body language	Extracts from *How to Communicate Successfully* (Wright, 1987, pp. 35–42), extracts from *Everybody's Guide to People watching* (Wolfgang, 1995, pp. 64–67), scenes from *New Headway Video, Beginner* (Murphy, 2002), facial expressions, eye contact, gestures, interpersonal space, touching, and so on, HW assignment—interview a foreigner about common gestures in their country
16. Negotiating plans in English and common discourse markers	"Weekend Away" activity from *Keep Talking* (Klippel, 1984, pp. 45–46), video of NSs planning a weekend away, natural ways to give opinions, agree and disagree in English, *could* and *would* modal auxilaries, *will* to confirm plans, present continuous to talk about fixed plans, discourse markers, role-play—planning a weekend away in Japan

5. Vocabulary test: The receptive vocabulary test used was Schmitt's Vocabulary Levels Test (Version 1) (Schmitt, 2000), which aims to test learners' receptive knowledge of 120 words from four different frequency ranges. According to Schmitt, Schmitt, and Clapham (2001), the Vocabulary Levels Test has good internal consistency, with a Cronbach alpha coefficient reported of over .9.
6. Discourse completion task (DCT): The DCT used was a Multimedia Elicitation Task (MET) developed by Gila Schauer (Lancaster University, UK). It consisted of 16 scenarios, with audiovisual prompts, requiring students to use eight different request speech acts with either same-status or higher status individuals. Students were provided with a Japanese translation of the audio prompts in order to avoid any listening comprehension problems and their oral responses were tape-recorded, transcribed, and blind-rated for pragmatic appropriateness by five native-speaking teachers, using rating guidelines (see Appendix S2 in online Supporting Information). The DCT's internal consistency (the degree to which individual items in the test "hang together") was investigated by calculating Cronbach's alpha coefficient. This was found to be .68 for the scale as a whole, which is just below the value of .7 considered the cutoff point for reliability. Analysis of the corrected item-total correlations indicated that scenarios 1, 2, 9, and 12 had particularly low values, below .3, which suggests that they were measuring something different from the scale as a whole. These were therefore removed from the final analysis, giving an improved Cronbach's alpha value of .8, which is considered reliable.
7. Oral interview: The oral interview used was based on a 1998–2000 version of the IELTS speaking test and consisted of five phases, lasting a total of 11–15 minutes. Interviews were conducted by 1 of 11 different native-speaker (NS) teachers (from the United States, the United Kingdom, Canada, or New Zealand) and recorded on both audiotape and videotape. The video interviews for both experimental and control groups were then blind-rated on five criteria (phonology, body language, fluency, context-appropriate vocabulary, and interactional competence) by three or four trained NS volunteers, using the descriptors provided (see Appendix S3 in online Supporting Information).
8. Student role-play: Once the results of the precourse testing had been analyzed, it became clear that the oral interview with a NS teacher was not giving students an opportunity to display their speaking skills in the best light. Many students appeared anxious meeting the NS teachers for the first time, which, to some extent, inhibited them in the interviews. Furthermore,

the IELTS oral test only gave students control over the conversation for 3 to 4 minutes, in the elicitation section (Phase 3), and this meant that they had limited opportunities to display their discourse competence (making topical moves, back-channeling, and so on). It was therefore decided to add a second speaking test in the form of a student role-play, which was administered at the beginning of the second semester and postcourse. Participants were asked to role-play a chance meeting with a friend in the street. After reading the role-play card, they were then given a few minutes of preparation time before performing the role-play in pairs. The role-plays were video-recorded and rated by the author on two criteria: (a) conversational behavior and (b) conversational management (see Appendix S4 in online Supporting Information). Eight postcourse role-plays, representing a total of 16 low-level and high-level students, were blind-rated by a second (trained) NS teacher and checked for interrater reliability using the Pearson product-moment correlation coefficient. There was a positive correlation between the mean scores on both components of the rating criteria. For conversational behavior, the Pearson correlation coefficient was slightly lower ($r = .70$, $n = 16$, $p = .003$), with a coefficient of determination (r^2) value of .49, indicating that the two variables shared 49% of their variance. Given the small sample size and the fact that an r value of .7 is viewed as large (Pallant 2005, p. 126), this scale was considered reasonably reliable. For conversational management, the Pearson correlation coefficient was higher ($r = .85$, $n = 16$, $p < .0005$), with a coefficient of determination (r^2) value of .72, indicating that the two variables shared 72% of their variance. This scale was also, therefore, considered reliable.

Table 2 summarizes how the communicative competence model was operationalized in the study.

Results

A one-way between-groups analysis of covariance (ANCOVA) was conducted to compare the effectiveness of the two different interventions designed to develop students' communicative competence. The independent variable was the type of intervention (textbook input or authentic input) and the dependent variables consisted of postcourse scores from the eight communicative competence measures. Participants' scores on the preintervention administration of the eight communicative competence measures were used as the covariates in this analysis. Preliminary checks were conducted to ensure that there was no

Table 2 Summary of communicative competence measures

Communicative competence component	Test
A. Linguistic competence	• Listening test • Pronunciation test • Grammar test • Vocabulary test • C-test • Oral interview (phonology and vocabulary sections)
B. Strategic competence	• Oral interview (interactional competence section) • Student role-play (conversational management section)
C. Pragmatic competence (pragmalinguistic + sociopragmatic)	• DCT • Oral interview (body language and context appropriate vocabulary use sections). • Student role-play (conversational behavior section)
D. Discourse competence	• Listening test • Oral interview (interactional competence and phonology sections) • Student role-play (conversational management section)

violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate.

After adjusting for preintervention scores, there were significant differences between the two intervention groups on postintervention scores for eight of the measures: (a) the listening component, $F(1, 59) = 4.44, p = .039, \eta_p^2 = .07$; (b) the receptive pronunciation component, $F(1, 58) = 11.84, p = .001, \eta_p^2 = .17$; (c) the receptive vocabulary component, $F(1, 58) = 14.81, p = .0005, \eta_p^2 = .20$; (d) the body language subcomponent of the IELTS oral interview, $F(1, 57) = 8.93, p = .004, \eta_p^2 = .14$; (e) the oral fluency subcomponent of the IELTS oral interview, $F(1, 57) = 5.01, p = .029, \eta_p^2 = .08$; (f) the interactional competence subcomponent of the IELTS oral interview, $F(1, 57) = 10.25, p = .002, \eta_p^2 = .15$; (g) the conversational behavior subcomponent of the student-student role-play, $F(1, 52) = 17.74, p < .0005, \eta_p^2 = .25$; and (h) the conversational management subcomponent of the student-student role-play, $F(1, 52) = 14.65, p < .0005, \eta_p^2 = .22$.

However, after adjusting for preintervention scores, there were no significant differences between the two intervention groups on postintervention scores for the remaining five measures: (a) the C-test component, $F(1, 58) = 2.69$,

$p = .11$, $\eta_p^2 = .04$; (b) the grammar component, $F(1, 59) = .022$, $p = .88$, $\eta_p^2 < 0005$; (c) the DCT component, $F(1, 54) = 1.73$, $p = .19$, $\eta_p^2 = .03$; (d) the pronunciation subcomponent of the IELTS oral interview, $F(1, 57) = 1.62$, $p = .21$, $\eta_p^2 = .03$; (e) the vocabulary subcomponent of the IELTS oral interview, $F(1, 57) = 2.02$, $p = .16$, $\eta_p^2 = .03$.

Table 3 summarizes precourse and postcourse mean scores and standard deviations for measures of communicative competence as a function of input condition.

Discussion

The results from the ANCOVA suggest strongly that after statistically controlling for differences in proficiency levels between participants, learners receiving the experimental treatment (authentic materials) developed their communicative competence to a greater degree than those receiving the control treatment (textbook materials). By performing separate statistical analyses on each of the various tests used in the trial, we are, to some extent, able to "tease apart" the individual components of communicative competence and investigate exactly how the two groups differed from each other after the intervention. The following discussion looks at each of these measures in turn.

Listening Test

The difference between the experimental and control groups in terms of listening proficiency shown in Table 3 was significant after the intervention ($F = 4.44$, $p = .039$), with a η_p^2 value of .07, indicating that 7% of the variance in postcourse scores could be accounted for by the treatment (a moderate effect size).

The quantity of listening input in both groups was similar over the 10-month study and, bearing in mind that a considerable proportion of the actual L2 listening practice students received in the classroom came from the same NS teacher, it is quite surprising to find any difference at all between the groups. There are two possible explanations for this. The first is that the increased focus on phonological aspects of English (such as stress-timing, weak forms, linking, and intonation units) in the experimental group succeeded in raising learners' awareness of these features of natural discourse and indirectly benefited their listening comprehension. The second possibility relates to the *quality* of the listening input learners were exposed to in the classroom. The textbook listening materials tended to have a slower speech rate and to display fewer features of natural, NS discourse and therefore did not prepare the students for the listening

Table 3 Precourse and postcourse mean scores and standard deviations for communicative competence measures as a function of input condition

	Precourse		Postcourse	
Source	*M*	*SD*	*M*	*SD*
Listening				
Experimental group	16.71	4.66	21.95	4.83
Control group	15.36	3.92	19.23	3.81
Receptive pronunciation				
Experimental group	45.83	4.19	48.58	3.49
Control group	44.48	2.99	45.42	3.60
C-Test				
Experimental group	84.43	12.29	95.10	9.83
Control group	78.97	9.75	89.71	7.70
Grammar				
Experimental group	106.97	10.99	109.32	7.78
Control group	100.13	12.53	106.87	8.27
Receptive vocabulary				
Experimental group	73.42	11.90	85.19	9.59
Control group	75.0	8.32	78.13	8.24
DCT				
Experimental group	2.86	0.40	3.05	0.31
Control group	2.77	0.30	2.87	0.37
IELTS (pronunciation)				
Experimental group	3.82	0.50	4.01	0.53
Control group	3.51	0.46	3.74	0.43
IELTS (body language)				
Experimental group	3.80	0.58	4.06	0.57
Control group	3.39	0.67	3.58	0.53
IELTS (fluency)				
Experimental group	3.64	0.61	3.84	0.62
Control group	3.28	0.57	3.39	0.51
IELTS (vocabulary)				
Experimental group	3.70	0.50	3.78	0.56
Control group	3.39	0.44	3.50	0.37
IELTS (interaction)				
Experimental group	3.52	0.71	3.75	0.72
Control group	3.25	0.78	3.24	0.56
Role-play (conv. behavior)				
Experimental group	4.17	0.80	4.48	0.54
Control group	3.26	0.53	3.45	0.55
Role-play (conv. management)				
Experimental group	4.06	0.92	4.40	0.62
Control group	3.26	0.51	3.47	0.55

test (which was natural-like) as well as the authentic materials. The most likely explanation is that the difference in listening comprehension between the two groups was due to a combination of both these factors, but this is speculation. Whatever the cause, the increased difficulty of the authentic listening materials used in the experimental group certainly did not appear to have a detrimental effect on learners' developing listening skills and this raises the question of how necessary it is to simplify listening texts for learners at intermediate to advanced levels of proficiency.

Receptive Pronunciation Test

The differences between the experimental and control groups in terms of their receptive comprehension of phonological features was significant after the intervention ($F = 11.84$, $p = .001$), with a η_2^p value of .17, indicating that 17% of the variance in postcourse scores could be accounted for by the treatment (a large effect size).

Because more time was spent focusing on phonology issues in the experimental class, it is hardly surprising to see this difference, and it supports the widely held belief that encouraging students to "notice" features of the target language can stimulate language acquisition. Most earlier investigations into the role of attention in L2 development have, however, focused principally on morphological or syntactical elements (Schmidt, 2001), so it is of interest to note that awareness-raising appears to facilitate the acquisition of other language features too. In addition to the increased focus on phonology in the experimental group, the authentic input probably gave students more opportunities to see these features of natural language put into practice because, as mentioned earlier, the contrived textbook listening materials often presented phonologically distorted samples of the L2 in an effort to ease the process of comprehension.

C-Test

No statistically significant differences were found between the groups in terms of their performance on the C-test ($F = 2.69$, $p = .11$). This is not particularly surprising, as this test focuses primarily on students' reading skills, whereas the intervention was designed predominantly to develop students' listening and speaking skills. Table 3 indicates that the mean scores on the C-test increased at a similar rate for both experimental and control groups (12.6% and 13.6%, respectively) and this can largely be attributed to the L2 texts and reading skills practice the learners experienced in other university classes over the period of the trial. This would appear to provide support to the view that the statistically

significant differences observed in other areas of the students' communicative competence were a direct result of the experimental intervention.

Grammar Test

No significant differences were found between groups in terms of their performance on the grammar test ($F = 0.022, p = .88$). Table 3 indicates that the mean scores on the grammar test increased only slightly for both experimental and control groups (2.2% and 6.7%, respectively) and, again, this is not surprising because students' grammatical competence was already quite well developed, after around 8 years of English instruction largely focused on grammatical aspects of the language. The control group received more grammar-focused input in their classes, as it was an integral part of the textbook syllabus, but, even so, it did not lead to any significant increases in their grammatical competence and this is probably because the grammatical items covered were often already familiar to the learners and therefore only served the function of reviewing old material. This highlights one of the problems for teachers using published coursebooks in Japan, which is that learners tend to be quite advanced grammatically but are not sufficiently prepared to cope with the listening or speaking materials associated with upper-intermediate or advanced textbooks produced for the international market. In other words, the Japanese education system, as it currently stands, does not produce learners who are balanced in terms of their communicative competence.

Receptive Vocabulary Test

The differences between the experimental and control groups in terms of their receptive comprehension of vocabulary was significant after the intervention ($F = 14.81$, $p < .0005$), with a η_p^2 value of .20, indicating that 20% of the variance in postcourse scores could be accounted for by the treatment (a large effect size).

The marked difference between the two groups is quite surprising, bearing in mind that the treatment condition only accounted for around 33% of the formal English input students received over the 10-month investigation period, and it suggests that the authentic materials were highly effective in developing learners' receptive comprehension of vocabulary. There are a number of possible reasons for this. First, the authentic materials exposed learners to richer input than the textbooks, with a greater number and wider range of vocabulary items, and therefore increased their chances of encountering and acquiring new words. Second, the predominant use of audiovisual materials in the experimental group meant that the new lexical material was highly contextualized when

it was presented to learners, and this is likely to have facilitated its acquisition. Third, because it is hypothesized that the authentic materials provided more motivating input than the textbooks were able to do, it could be that learners engaged more in the learning experience in the experimental group (see Peacock, 1997). This is certainly the impression conveyed in students' classroom diaries, in which they recorded their perceptions of what each class "was about" and their feelings on the content.[6] Schumann (1997) also provided an interesting perspective on this in his stimulus appraisal model, which sees language input as being evaluated along five criteria by learners: novelty, intrinsic pleasantness, goal/need significance, coping mechanisms, and self/social image. Input that is evaluated positively against these criteria is believed to result in greater engagement and sustained deep learning, whereas negative appraisals result in avoidance. In this investigation, the authentic input is more likely to have been evaluated positively, when compared with the textbook input, on at least three of the criteria: (a) novelty (both the authentic materials themselves and their accompanying tasks were new to students), (b) pleasantness (the teacher's knowledge of the students' interests allowed the selection of materials likely to be considered enjoyable), and (c) goal/need significance (again, the teacher's intimate knowledge of the groups' language needs and future goals facilitated selection of appropriate materials). For the remaining two of Schumann's criteria, the effect of authentic input is less predictable. In terms of coping mechanisms, we might expect the challenging nature of the experimental materials to have evoked a negative appraisal, but this was probably avoided by careful task design, which ensured that learners were able to experience success in the activities they were asked to perform. Effective learning (and positive evaluation) is most likely to be achieved when students have appropriate levels of both challenge and support in the classroom (Gilmore, 2009; Mariani, 1997)—a condition realized in the experimental treatment by varying the *task* rather than the *text*. Finally, the self/social image criterion, which assesses the compatibility of an experience against social or cultural norms, could have resulted in a negative evaluation because the experimental materials were so different from those used in a traditional Japanese ELT classroom. Perhaps this did not happen in this case because of positive evaluations on the other criteria.

Discourse Completion Task

No significant differences were found between groups in terms of their performance on the discourse completion task ($F = 1.73$, $p = .19$). This was disappointing, and unexpected, because the experimental syllabus included materials

and tasks specifically designed to develop learners' pragmatic competence (see Table 1). There are two possible explanations for this lack of difference:

1. The first is that the DCT was simply not sensitive enough to detect changes in the learners' pragmatic competence. It only focused on the use of a single speech act (requests) rather than a range of different speech acts. Interlanguage pragmatics studies to date have found no speech communities that lack speech acts for requesting, suggesting, inviting, refusing, apologizing, complaining, complimenting, and thanking (Kasper & Schmidt, 1996), so perhaps these could act as a basis for a more sensitive DCT. In addition, the NS raters were only given written transcriptions of learners' utterances rather than the taped responses themselves, which meant that all the paralinguistic details, such as tone of voice, pitch changes, or loudness (which are an important component of affective speech), were lost and learners were rated only on the actual words they used. Although the NS raters received quite detailed instructions and guidance on the rating procedure (see Appendix S2 in online Supporting Information), their scores for the pragmatic appropriateness of students' responses differed markedly at times. An estimate of the DCT's reliability was obtained by comparing the author's postcourse ratings with those of all other NS raters, using the Pearson product-moment correlation coefficient. There was a positive correlation between the mean scores ($r = .60$, $n = 59$, $p < .0005$), but a coefficient of determination (r^2) value of .36 means that the two variables shared only 36% of their variance, suggesting the DCT had low levels of reliability. As a further estimate of the test's reliability, a NS's responses to the DCT were added in for the postcourse rating as a control measure (the NS raters were not aware of this because all responses were rated blind). The raters *did* rate the NS's responses as the most appropriate, giving her a mean score of 4.35, well above the combined total mean of 3.10 for the students in the treatment groups. However, they sometimes varied dramatically in their ratings of the NS's responses, in some scenarios giving a score between 1 and 5 for the same utterance. Perhaps a DCT design using multiple-choice responses, with only one pragmatically appropriate answer and four nonnative-speaker (NNS) distracters, would result in a more reliable measure as well as being much faster to implement.
2. The second possible explanation is that the pragmatics training implemented in the experimental group did not lead to the desired increase in students' pragmatic competence. Although the materials used in the experimental treatment did look at issues of register (see Table 1), request speech

acts were not dealt with specifically and it could be that development of pragmatic competence relies more on the memorization of specific fixed phrases for particular contexts than on real-time construction of utterances. If this were the case, more general advice in the classroom on producing pragmatically appropriate English might lead to higher levels of awareness of pragmatic issues, but it would not, in itself, necessarily translate into better performances on a DCT.

Finally, it should be pointed out that although the results from the DCT showed no statistically significant differences between the treatment groups, subscores on the IELTS oral interview and student role-play also measured pragmatic aspects and there is evidence from these measures that some aspects of the pragmatics training, such as opening and closing conversations, were effective.

IELTS Oral Interview

The differences between the experimental and control groups for the combined means of all five components of the IELTS oral interview were significant after the intervention ($F = 6.84, p = .011$), with a η_p^2 value of .11, indicating that 11% of the variance in postcourse scores could be accounted for by the treatment (a moderate effect size). However, the statistical analysis of the individual components of the IELTS test shown next gives a clearer picture of changes in the learners' communicative competence over the 10-month trial.

Pronunciation Component of the IELTS Oral Interview

No statistically significant differences were found between the groups in terms of their performance on the pronunciation component of the IELTS oral interview ($F = 1.62$, $p = .21$). Table 3 shows that the pronunciation mean scores for both groups in the IELTS test did increase slightly over the trial period and this was to be expected because there was some pronunciation focus in both syllabi. The experimental group received more explicit training on phonological features of English, and this appears to have had a larger effect on their comprehension (with a highly significant difference seen in the receptive pronunciation test) than their production. Perhaps improvements in productive pronunciation require more intensive periods of training than the 9 hours or so learners in the experimental group received or longer periods of consolidation than the 10-month duration of the trial. Additionally, the fact that the learners in the study were all at a relatively advanced level, with around 8 years of L2 instruction at school and university behind them, means that they may have been less likely to change their phonological behavior.

Body Language Component of the IELTS Oral Interview

The difference between the experimental and control groups for the body language component of the IELTS oral interview was significant after the intervention ($F = 8.93$, $p = .004$), with a η_p^2 value of .14, indicating that 14% of the variance in postcourse scores could be accounted for by the treatment (a large effect size). This suggests that the sociopragmatic training (on facial expressions, gestures, eye contact, and proxemics) implemented in the experimental group did successfully encourage learners to alter their behavior toward the norms of the target speech community. This is an important result because little empirical evidence exists (at least in the field of applied linguistics) for the success of this kind of training (although see Collett, 1971, for a notable exception). Nonverbal communication (NVC) is widely recognized as being crucial for successful communication, and when people from very different cultures interact, sociopragmatic misunderstandings are frequent and often serious (Argyle, 1988; Bailey, 1997; Bilbow, 1997; Chick 1985, 1989; Gumperz, 1982; Wolfson, 1989), yet this area is rarely given much space in current foreign language teaching syllabi. If we want to take the notion of communicative competence seriously in our profession, then it is crucial that we begin to broaden our training programs to include all of its different dimensions, not just those with which we are most familiar. The results presented here suggest that NVC training in the classroom can improve learners' sociopragmatic competence and I would argue for its inclusion when there are wide disparities between the students' culture and the target culture.

Fluency Component of the IELTS Oral Interview

The difference between the experimental and control groups for the fluency component of the IELTS oral interview was significant after the intervention ($F = 5.01$, $p = .029$), with a η_p^2 value of .08, indicating that 8% of the variance in postcourse scores could be accounted for by the treatment (a moderate effect size). Table 3 shows that both the experimental and control groups improved in fluency over the 10-month investigation and this was anticipated because both groups were given numerous speaking opportunities in class: Like any other skill, speaking improves with practice. What is of particular interest here, though, is what led to the significant difference in fluency between the two groups. Because both treatments involved similar quantities of student talking time, the enhanced fluency in the experimental group is likely to have come from the explicit focus on conversational strategies (see Table 1). Perhaps the learners were using hesitation devices more often or more effectively, thus appearing more fluent or perhaps the focus on discourse intonation and tone groups

encouraged learners to pause in more appropriate places in the discourse (e.g., at transition relevance places). Without a quantitative analysis of the interview transcriptions, which is outside the scope of this investigation, it is impossible to come to any firm conclusions on this matter. What seems clear, however, is that some aspect of the explicit focus on conversational strategies had a beneficial effect on learners' fluency and this supports the use of awareness-raising, or "noticing," strategies in the classroom.

Appropriate Vocabulary Use Component of the IELTS Oral Interview

No statistically significant differences were found between the groups in terms of their performance on the vocabulary component of the IELTS oral interview ($F = 2.02$, $p = .16$). Examination of Table 3 suggests that the mean scores for both groups changed very little over the period of the investigation on this criterion and there are two possible explanations for this. The first is that there were changes in the appropriateness of students' vocabulary use but that these were difficult for raters to pick up while watching the video and, at the same time, grading all five subcomponents of the IELTS exam. This is a distinct possibility because, as we saw earlier, differences between the control and experimental groups on receptive vocabulary use were statistically significant at the end of the trial. Furthermore, rating appropriate vocabulary use is extremely difficult in real time and, without the benefit of a written transcript, decisions on the appropriateness of each individual word have to be made instantaneously. It is likely that this imposes unrealistic demands on the raters. The second possible explanation is that the learners' productive use of new vocabulary lagged behind their receptive comprehension. This is a possibility because we might expect learners to require numerous encounters with new words before they feel sufficiently confident to begin using them in their own discourse. Perhaps the 10-month trial period was simply too short to detect these kinds of changes.

Interactional Competence Component of the IELTS Oral Interview

The difference between the experimental and control groups for the interactional competence component of the IELTS oral interview was significant after the intervention ($F = 10.25$, $p = .002$), with a η_p^2 value of .15, indicating that 15% of the variance in postcourse scores could be accounted for by the treatment (a large effect size). The mean scores in Table 3 show that the control group did not change at all on this measure, whereas the experimental group improved 6.5%. Again, this suggests that the explicit focus on conversational strategies (such as turn-taking; developing conversation; using

reactive tokens, discourse markers, or hesitation devices) in the experimental treatment was effective in bringing about changes in the learners' strategic or discourse competencies. This result also has potentially important implications for language teaching. Insights into the discourse strategies employed by NSs during conversation have only recently become available to teachers, thanks to the rapidly expanding fields of discourse and conversational analysis. Although some language textbooks and resource books have begun incorporating these insights into their designs (see, e.g., Dörnyei & Thurrell, 1992; Nolasco & Arthur, 1987), little empirical research currently exists to support the awareness-raising of discourse features in the classroom. Intuitively, it makes sense that they would benefit learners: The recognized value of "noticing" on acquisition probably applies to all features of language, not just grammatical items, which are generally the focus of attention. The results presented here therefore support the incorporation of training in conversational strategies into the classroom.

Student Role-Play

The differences between the experimental and control groups for both the conversational behavior component and the conversational management component of the student role-play were significant after the intervention for (a) the conversational behavior scale ($F = 17.74$, $p < .0005$), with a η_p^2 value of .25, indicating that 25% of the variance in post-course scores could be accounted for by the treatment (a large effect size) and (b) the conversational management scale ($F = 14.65$, $p < .0005$), with a η_p^2 value of .22, indicating that 22% of the variance in postcourse scores could be accounted for by the treatment (also a large effect size). This pronounced difference between the two treatment groups was seen despite the fact that the period of investigation was half that of the other communicative competence measures (the first role-play was conducted in September 2004 rather than April 2004). The conversational behavior measure focused on sociopragmatic aspects of the students' performance (see Appendix S4 in online Supporting Information) and was therefore similar to the body language component of the IELTS oral interview, except that the interaction was NNS-NNS, rather than NS-NNS. The results provide further support that NVC training can produce a marked change in student behavior, toward NS norms, in relatively short periods of time. The difference in performance between the two groups was largely anticipated because the experimental treatment involved specific advice on and practice with the features measured by this rating descriptor, but it is encouraging to see that the learners were so readily able to incorporate these behaviors into their own

productive repertoires. The conversational management component of the role-play focused principally on discoursal aspects of the students' performance (see Appendix S4 in online Supporting Information) and was therefore similar to the interactional competence component of the IELTS oral interview. The role-play was, however, considered a more sensitive measure of students' discourse competence because learners had more opportunities to take responsibility for topical coherence and topic development in the absence of a NS interlocutor. The results for this measure strongly support the earlier conclusion that an explicit focus on conversational strategies can benefit learners' strategic or discourse competencies. Whereas the control group was only given opportunities to "do speaking" in pair or group-work activities, the experimental group received specific awareness-raising of conversational strategies in English as well as practice incorporating these strategies into their own conversations. The significant difference between the two groups in terms of their role-play performances therefore supports a language teaching methodology that aims at explicit awareness-raising and practice of discourse features rather than one that simply provides students with speaking opportunities, as traditional conversation classes tend to do.

Conclusion

The results of this study strongly suggest that the authentic materials used with the experimental group in the investigation were better able to develop a range of communicative competencies in the learners than the two EFL textbooks used with the control group. This finding was predicted on the grounds that the authentic materials, with their associated tasks and activities, provided richer input for learners to work with in the classroom, which, in turn, allowed them to notice and then acquire a wider variety of linguistic, pragmatic, strategic, and discourse features. The consciousness-raising was therefore facilitated by (a) providing participants with rich input and (b) drawing learners' attention to useful features through careful task design and follow-up practice activities.

All of the components of the communicative competence model appear to have been amenable to training to some degree, when the classroom context allowed learners to attend to, or notice, relevant linguistic or paralinguistic features in the input. The extent of development of each component is likely to have been affected by two factors: (a) the preintervention levels of competence existing in the student population investigated and (b) the types of input or tasks selected for use in the classroom. The same experimental treatment would probably lead to very different results with a group of learners from another

culture; for example, it is unlikely that statistically significant differences in sociopragmatic measures would have been observed in European students, for which behavioral norms are closer to those of NSs of English. Similarly, the lack of statistically significant results for the grammar test may have been due to the already well-developed grammatical competence exhibited by the Japanese subjects.

The least successful aspect of the experimental treatment was associated with the development of students' pragmalinguistic competence. Here, a general focus on issues of register did not noticeably affect learners' performance on the DCT, although other pragmatic measures employed did detect statistically significant differences between the experimental and control groups. It could be that students require specific training on speech acts for specific contexts in order to generate observable changes; this was seen in the experimental group, for which a focus on strategies for closing down conversation resulted in improved pragmatic competence in the student role-play. The DCT itself proved to be a rather unreliable measure of pragmalinguistic competence, and much more work needs to go into producing a testing instrument that is both reliable and practical to implement in the classroom.

If the different components of the communicative competence model are amenable to training, as seems to be the case, it begs the question: Why do the majority of EFL textbooks continue to focus predominantly on lexicogrammatical features? A number of possible reasons exist for the current status quo. First, teaching form is well established and safe (Thornbury, 1999) and publishers are reluctant to take risks with innovative materials, given the enormous costs of developing textbooks for the international market (Tomlinson, 2001). Second, our understanding of linguistic competence (vocabulary, syntax, phonology) is far more comprehensive than it is for the other components of the model, and its elements are currently more amenable to itemization and testing.

One possible practical response to these issues would be for materials designers to abandon attempts to organize content around a structural syllabus (particularly at higher proficiency levels) and to instead provide learners with rich samples of authentic input. The communicative competence model could be used to inform the syllabus, ensuring that learners' linguistic, strategic, pragmatic, and discourse competencies were all developed appropriately. This is precisely what the study reported on here attempted to do and, as we have seen, the results appear to be very promising.

However, a number of difficulties stand in the way of implementing this kind of syllabus. First, teachers do not always have access to a wide range of authentic materials, and even when they do, they often lack the necessary

expertise to exploit them to effectively meet their students' needs. NNS teachers are particularly disadvantaged, as it is naturally more difficult for them to judge the pragmalinguistic or sociopragmatic appropriateness of materials. Second, communicative competence-centered approaches, using authentic materials, can be extremely time-consuming to implement. Initially, they require some kind of needs analysis to determine how best to develop learners' overall competence, and decisions need to be taken in terms of how much emphasis to place on the different components of the model. After this, appropriate materials need to be collected and effective tasks designed to highlight or practice the discourse features of interest. The final difficulty associated with this kind of approach is the design of tests, which are both a fair reflection of the course content and practical to implement. Because the syllabus is not preconceived, but is rather co-constructed by participants during the course, what takes place in the classroom cannot be predicted beforehand and must be examined retrospectively. Fair assessment can therefore only occur if tests are designed specifically for each course. In addition, any syllabus that aims to develop a broad range of communicative competencies in learners must also endeavor to test them. If we continue to assess only those features of the language that are easily measured (often lexicogrammatical items), the "backwash effect" will ensure that students and teachers remain firmly focused on them at the expense of other areas. This remains problematic because reliable measures of strategic, pragmalinguistic, sociopragmatic, and discourse competencies have not been established yet. Strategic competence is difficult to assess because instances of communication breakdown are usually infrequent and unpredictable in conversation, so tests would probably need to involve some kind of elicitation. Pragmalinguistic competence is also extremely difficult to assess; DCTs are very time-consuming to implement and NSs often disagree on the level of appropriateness of students' responses when judging pragmatic rather than grammatical features. Measuring sociopragmatic competence normally involves analyzing students' behavior in a particular context, and role-play scenarios are probably the best method for assessing this area. Discourse competence is difficult to assess because it requires analysis of longer stretches of spoken or written discourse. With writing samples, this means focusing on the overall structure of a piece of work and assessing how cohesive/coherent it is or how well it approximates the generic model. With spoken samples, it means analyzing conversational turns for cohesion or coherence or identifying whether longer turns (e.g., in oral narratives) include all of the obligatory parts. This is very difficult to do accurately without transcripts of the conversation (something which is, of course, impractical in the classroom) and also requires

a high degree of language awareness from the assessor. Despite the difficulties in assessing learners' overall communicative competence outlined here, this is an issue that needs to be engaged with and discussed so that practical solutions can be found.

In conclusion, expanding the curriculum to include all aspects of the communicative competence model raises some difficult questions: How can we develop a broader range of competencies in our learners without increasing the time allocated to language courses? Which components of the model should we focus on with specific groups of English language learners and how might the balance change as their interlanguage develops? How can we accurately assess the communicative competence of our learners, and are the methods available to us practical or realistic to implement in the classroom? Space here does not allow further discussion of these issues, but it seems clear that the communicative competence model and its pedagogical implications are areas worthy of renewed interest by the applied linguistics community.

Revised version accepted 14 January 2010

Notes

1 The construct described here has, however, been criticized more recently because it models itself on educated native speakers and takes their communicative competence as the ultimate goal of foreign language learning. This, it is argued, devalues the social identity and competencies that learners have already developed in their own culture and neglects the fact that the communicative needs of nonnative speakers of English are very different from those of native speakers, operating in specific speech communities. Byram and Fleming (1998) instead proposed a model based around intercultural communicative competence, which emphasizes the knowledge and skills needed to understand and to successfully communicate with, people from other, unfamiliar cultures. In this sense, it extends and builds on the communicative competence model proposed here rather than abandoning it altogether.

2 "Authenticity" is an extremely difficult concept to define, with at least eight, interrelated meanings emerging from the literature (for a summary, see Gilmore, 2007b, p. 98).

3 The same teacher was used for both experimental and control groups in order to avoid introducing further uncontrolled variables into the investigation. It could be argued that the researcher/teacher's greater investment in the experimental materials was a threat to the validity of the results, but every effort was taken to ensure that both groups in the trial were taught with equal commitment and enthusiasm.

4 With the exception of the student role-play, which was administered midcourse, in September 2004, and postcourse, in January 2005.
5 Based on the authors' acknowledgements of copyright material used in the two textbooks, none of the input in *Face to Face* and only five texts in *Inside English* (including the lyrics for two songs, extracts from two books, and an extract from a Web page) were authentic.
6 The qualitative results from this study will be reported on in a later publication.

References

Alderson, J., & Beretta, A. (Eds.). (1992). *Evaluating second language education*. Cambridge: Cambridge University Press.

Allen, E., Bernhardt, E. B., Berry, M. T., & Demel, M. (1988). Comprehension and text genre: An analysis of secondary foreign language readers. *Modern Language Journal, 72*, 63–72.

Allwright, D., & Bailey, K. (1991). *Classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.

Argyle, M. (1988). *Bodily communication*. London: Routledge.

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

Bachman, L., & Palmer, A. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly, 16*, 449–465.

Bacon, S., & Finnemann, M. (1990). A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *Modern Language Journal, 74*, 459–473.

Bailey, B. (1997). Communication of respect in interethnic service encounters. *Language in Society,* 26, 327–356.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*, 233–262.

Batstone, R. (1996). Noticing. *ELT Journal, 50*, 273.

Biber, D., Conrad, S., & Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics, 15*, 169–189.

Bilbow, G. T. (1997). Cross-cultural impression management in the multicultural workplace: The special case of Hong Kong. *Journal of Pragmatics, 28*, 461–487.

Bradford, B. (1988). *Intonation in context*. Cambridge: Cambridge University Press.

Brazil, D., Coulthard, M., & Johns, C. (1980). *Discourse intonation and language teaching*. London: Longman.

Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bruner, J. (1983). *Child's talk: Learning to use language*. New York: Norton.
Bublitz, W. (1988). *Supportive fellow speakers and cooperative conversations*. Amsterdam: Benjamins.
Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon, UK: Multilingual Matters.
Byram, M., & Fleming, M. (Eds.) (1998). *Language learning in intercultural perspective*. Cambridge: Cambridge University Press.
Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2–27). London: Longman.
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.
Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics,* 6(2), 5–35.
Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in Communicative Language Teaching? *TESOL Quarterly, 31*, 141–152.
Chick, K. (1985). The interactional accomplishment of discrimination in South Africa. *Language in Society, 14*, 299–326.
Chick, K. (1989). Intercultural miscommunication as a source of friction in the workplace and in educational settings in South Africa. In O. Garcia & R. Otheguy (Eds.), *English across cultures, cultures across English* (pp. 139–160). New York: Mouton de Gruyter.
Clancy, P. A., Thompson, A., Suzuki, R., & Tao H. (1996). The conversational use of reactive tokens in English, Japanese and Mandarin. *Journal of Pragmatics, 26*, 355–387.
Clyne, M. (1979). Communicative competencies in contact. *ITL, 43*, 17–37.
Collet, P. (1971). Training Englishmen in the non-verbal behaviour of Arabs: An experiment on intercultural communication. *International Journal of Psychology, 6*, 209–215.
Crow, B. (1983). Topic shifts in couples' conversations. In R. Craig & K. Tracey (Eds.), *Conversational coherence: Form, structure and strategy* (pp. 136–156). Beverly Hills, CA: Sage.
Day, R. (Ed.). (1986). *Talking to learn: Conversation in second language acquisition*. Rowley, MA: Newbury House.
Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing, 9*, 187–206.
Dörnyei, Z., & Thurrell, S. (1992). *Conversation and dialogues in action*. Hemel Hempstead, UK: Prentice Hall.
Eggins, S., & Slade, D. (1997). *Analysing casual conversation*. London: Cassell.

Ellis, G., & Sinclair, B. (1989). *Learning to learn English*. Cambridge: Cambridge University Press.

Falla, T. (1994). *Headway video* (Pre-Intermediate). Oxford: Oxford University Press.

Freeman, D., & Holden, S. (1986). Authentic listening materials. In S. Holden (Ed.), *Techniques of teaching* (pp. 67–69). London: Modern English Publications.

Fuller, D., & Fuller, C. (1999). *Face to face*. Tokyo: Macmillan.

Gilmore, A. (2007a). *Getting real in the language classroom: Developing Japanese students' communicative competence with authentic materials*. Unpublished doctoral dissertation, Nottingham University, UK.

Gilmore, A. (2007b). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*, 97–118.

Gilmore, A. (2009). The times they are a-changin': Strategies for exploiting authentic materials in the language classroom. In S. Rilling & M. Dantas-Whitney (Eds.), *Authenticity in adult classrooms and beyond* (pp. 155–168). Alexandria, VA: TESOL.

González, O. (1990). *Teaching language and culture with authentic materials*. Unpublished doctoral dissertation, West Virginia University, Morgantown.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: Vol. 9*, *Pragmatics* (pp. 41–58). New York: Academic Press.

Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.

Halliday, M., & Hassan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.

Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 401–435). Rowley, MA: Newbury House.

Hopkins, D., & Nettle, M. (1995). *Passport to IELTS*. London: Prentice Hall.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Socio-linguistics: Selected readings* (pp. 269–293). Harmondsworth, UK: Penguin Books.

Johnson, K. (1995). *Understanding communication in second language classrooms*. Cambridge: Cambridge University Press.

Jones, R. (2001). A consciousness-raising approach to the teaching of conversational storytelling skills. *ELT Journal, 55*(2), 155–163.

Kasper, G. (2001a). Four perspectives on L2 pragmatic development. *Applied Linguistics, 22*, 502–530.

Kasper, G. (2001b). Learning pragmatics in the L2 classroom. In L. F. Bouton (Ed.), *Pragmatics and language learning monograph series, Vol. 10* (pp. 1–22). Urbana: Division of English as an International Language, University of Illinois at Urbana-Champaign.

Kasper, G., & Rose, R. K. (1999). Pragmatics and SLA. *Annual Review of Applied Linguistics, 19*, 81–104.

Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition, 18*, 149–169.

Kienbaum, B. E., Russell, A. J., & Welty, S. (1986). *Communicative competence in foreign language learning with authentic materials: Final project report*. Washington: Center for International Education, Purdue University.

Klein-Braley, C., & Raatz, E. (1984). A survey on the C test. *Language Testing, 1*, 134–146.

Klippel, F. (1984). *Keep talking: Communicative fluency activities for language teaching*. Cambridge: Cambridge University Press.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Leech, G. N. (1983). *Principles of pragmatics*. London: Longman.

Little, D., Devitt, S., & Singleton, D. (1989). *Learning foreign languages from authentic texts: Theory and practice*. Dublin: Authentik in association with CILT.

Little, D., & Singleton, D. (1991). Authentic texts, pedagogical grammar and language awareness in foreign language learning. In C. James & P. Garrett (Eds.), *Language awareness in the classroom* (pp. 123–132). London: Longman.

Maggs, P., Kay, S., Jones, V., & Kerr, P. (2004). *Inside English*. Oxford: Macmillan.

Mariani, L. (1997). Teacher support and teacher challenge in promoting learner autonomy. *Perspectives, 23*(2), 1–10.

McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.

McCarthy, M., & Carter, R. (1994). *Language as discourse*. Harlow, UK: Longman.

McGarry, D. (1995). *Learner autonomy 4: The role of authentic texts*. Dublin: Authentik.

Morrison, B. (1989). Using news broadcasts for authentic listening comprehension. *ELT Journal, 43*(1), 14–18.

Morrow, K. (1977). Authentic texts and ESP. In S. Holden (Ed.), *English for specific purposes* (pp. 13–17). London: Modern English Publications.

Murphy, J. (2002). *New Headway video* (Elementary). Oxford: Oxford University Press.

Murphy, R. (1994). *English grammar in use*: *Intermediate level* (2nd ed.). Cambridge: Cambridge University Press.

Naunton, J. (1994). *Think first certificate*. Harlow, UK: Longman.

Nolasco, R., & Arthur, L. (1987). *Conversation*. Oxford: Oxford University Press.

Nunan, D. (1991). *Language teaching methodology*. Hemel Hempstead, UK: Prentice Hall.

Nunan, D. (1996). Issues in second language acquisition research: Examining substance and procedure. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 349–374). San Diego, CA: Academic Press.

Pallant, J. (2005). *SPSS survival manual* (2nd ed.). Maidenhead, UK: Open University Press.

Peacock, M. (1997). The effect of authentic materials on the motivation of EFL learners. *ELT Journal, 51*(2), 144–156.
Ritchie, W., & Bhatia, T. (Eds.) (1996). *Handbook of second language acquisition*. San Diego, CA: Academic Press.
Rogerson, P., & Gilbert, J. B. (1990). *Speaking clearly*. Cambridge: Cambridge University Press.
Sakamoto, N., & Naotsuka, R. (1982). *Polite fictions*. Tokyo: Kinseido.
Salsbury, T., & Bardovi-Harlig, K. (2001). "I know your mean, but I don't think so": Disagreements in L2 English. *Pragmatics and Language Learning, 10*, 131–151.
Schauer, G. A. (2009). *Interlanguage pragmatic development: The study abroad context*. London: Continuum.
Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137–174). Rowley, MA: Newbury House.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–152.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
Schmidt, R., & Richards, J. (1980). Speech acts and second language learning. *Applied Linguistics, 1*, 127–157.
Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing, 18*, 55–88.
Schumann, J. H. (1997). *The neurobiology of affect in language*. Oxford: Blackwell.
Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Soars, J., & Soars, L. (1987). *Headway upper-intermediate student's book*. Oxford: Oxford University Press.
Soars, J., & Soars, L. (1993). *Headway elementary student's book*. Oxford: Oxford University Press.
Soars, J., & Soars, L. (1996). *New Headway English course intermediate student's book*. Oxford: Oxford University Press.
Soars, J., & Soars, L. (2000). *New Headway English course pre-intermediate student's book*. Oxford: Oxford University Press.
Swaffar, J. (1985). Reading authentic texts in a foreign language: A cognitive model. *Modern Language Journal, 69*, 15–34.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*, 91–109.
Thornbury, S. (1999). Window-dressing vs cross-dressing in the EFL sub-culture. *Folio, 5*(2), 15–17.
Tomlinson, B. (2001). Materials development. In R. Carter & D. Nunan (Eds.), *Teaching English to speakers of other languages* (pp. 66–71). Cambridge: Cambridge University Press.
Verne, J. (1873). *Around the world in eighty days*. London: Sampson Low, Marston Low & Searle.
Wigglesworth, G. (2005). Current approaches to researching second language learner processes. *Annual Review of Applied Linguistics, 25*, 98–111.
Wolfson, N. (1989). *Perspectives*. Boston: Heinle & Heinle.
Wolfgang, A. (1995). *Everybody's guide to people watching*. Yarmouth, ME: Intercultural Press.
Wright, A. (1987). *How to communicate successfully*. Cambridge: Cambridge University Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1: C-Test.
Appendix S2: DCT Rating Guidelines and Sample Scenarios (see also Schauer 2009, pp. 216–233).
Appendix S3: Rating Descriptors for IELTS Oral Interview.
Appendix S4: Rating Descriptors for Student Role-Play.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.