Lexical Density of Invented Dialogues in EFL Textbooks: How Natural Are They as Models of Spoken Discourse?[1]

TSURII, Chie

1. Introduction

The main purpose of this paper centres on examining how natural invented dialogues appearing in EFL textbooks are. I will analyse texts which are used as model dialogues in EFL textbooks aimed at improving the speaking ability of pupils in senior high schools in Japan. The texts will be analysed in terms of lexical density.

In Japan, guidelines for teaching by the Ministry of Education were promulgated in 1989, and from 1994, new English language related curriculum components known as 'oral communication A', 'oral communication B', and 'oral communication C' have been introduced. The purpose of these modules is, according to the government guidelines (Monbusho, 1989 : 108), "to improve pupils' attitudes towards attempting to communicate in English". The focus on communicating in English learning seems to have been taken very seriously.[2]

Most of the present students at university received their English language education based on the above curriculum. However, many of my students do not seem to have improved in their attitude towards communication, nor in their ability to communicate, compared with students several years ago. On the contrary, some of them seem to suffer from an inferiority complex regarding their inability to speak English fluently and accurately.

I teach several courses in listening and speaking ability to first-year

students at university. My students major in either sociology, business administration, or economics. Many of them often say that they do not think they are good at English, especially listening and speaking skills. Some students can not speak English at all despite the fact that they have a grammatical knowledge of the language, and when I listen to them attempt to practice, they often say that they have nothing to say or do not know how they can express themselves in English. During the lesson I often encounter situations where students who are asked a question answer just 'Yes' or 'No', or answer in one short sentence, and then remain silent waiting for the next question.

It seems that one cause of this lies in the use of invented dialogues used as models in EFL textbooks. Although students may have the opportunity to hear authentic conversation inside and outside the classroom, their English input is provided mainly by textbooks. English is used as a foreign language, not a second language, in Japan, therefore we seldom use English in everyday life. In addition to this, for most senior high school pupils, the authentic use of English used in, for example, cinematic dialogues and in news programmes, is demanding. In 'oral communication' classes, pupils listen to and practice reading the dialogue with instructions on the expressions used in the dialogue given by the teacher. In some cases, pupils may memorise the dialogue and may act out the dialogue with a partner in front of the class and the teacher. For these pupils, dialogues in textbooks are essential and may be influential in terms of the pupils' progress.

In order to investigate the differences between invented dialogues and authentic conversations, I will analyse dialogues in EFL textbooks in terms of their lexical density. First of all, I will summarise the differences between spoken and written language and lexical density; and then, analyse the texts appearing in EFL textbooks used in Japanese senior high schools in terms of their lexical density, in order to investigate the naturalness of the dialogues as models of spoken language. I will see whether the lexical density of the invented dialogues is higher or lower than that of the authentic dialogues. It is possible that invented dialogues will have higher density because of deficiencies in the charac-

teristics which are peculiar to spoken language.

2. Speech versus Writing: The Difference between Spoken and Written Language

In using the English language, there are a great many different kinds of discourse: telephone conversations, face-to-face conversations, public conversations, prepared speeches, broadcasts, personal letters, general fiction, press editorials, academic prose, official documents, romance fiction, biographies, press reportage, and so on. (These terms of categorisation are taken from Biber, Conrad and Reppen 1998 : 152-4).

If we consider the differences among a variety of language uses in terms of the medium of production, these are traditionally divided into two groups—spoken and written—although, as will be shown below, current thinking blurs the distinction between them. In spoken activities, we use mainly our mouth and ears, the speaker producing the sound and the other participant(s) receiving and decoding the sound. On the other hand, in written activities we use our hands and eyes, using the letters to convey information to the reader. We can say that the first five categories I list above are types of spoken discourse, and the last eight are types of written discourse.

In order to clarify the distinctions between spoken and written language, it is also necessary to take into account both the contextual difference between both types of discourse and their formal linguistic characteristics. Halliday (1985 : 78) states in this respect that "[o]ne has to think of both written and spoken language in terms of three interrelated aspects: the nature of the medium, the functions served, and formal properties displayed—let us say function, medium, and form".

One of the differences in production context is this: written discourse is permanent, while, spoken discourse is transient. Stubbs (1996 : 72) states in this regard that "[o]n average, a written text is . . . permanent, highly edited, redrafted and rehearsed, rather than being unplanned and spontaneous as most casual conversation is". In written discourse, writers have time to plan, that is, to organise their thoughts in written

form, to elaborate the wording, the structure of the text, and the logic behind it, and to check and edit the text, in order to convey ideas adequately. Readers also have time to read at their own pace, and re-read, while consulting a dictionary. They also have the opportunity to come to terms with the subject matter. By contrast, in spoken discourses, speakers often do not have enough time to think and pause. Turns can be taken by another interlocutor (s) during the speaking activity; some clues such as the syntactic completeness, pause, and drop in the vocal pitch can be seen as indicating the end of the current speaker's turn at speaking; then listeners are always attentive to these clues (see McCarthy, 1991 : 104, 127). The listener also does not have enough time to think over what the speaker is talking about. We need to understand what the speaker says while simultaneously replying in an adequate manner to the changing context. We can also say that writing is a slower activity, and speaking is a faster activity.

The other difference in context between spoken and written language is this: in spoken activities, participants share the same context; in writing on the other hand, writers and readers do not often share the same situation. In other words, communication in spoken discourse is 'two-way', and in written discourse is 'one-way'. As Nunan (1993 : 8) states, "written language is used to communicate with others who are removed in time and space". To echo Tribble (1996 : 10), "[t]he writer has never met and will probably never meet the person who processes their request". Writers cannot count on the readers' reaction or feedback, or on other contextual cues. In the case of speaking activities, both the speaker and the listener can rely on the contextual clues since communications often take place face-to-face. The speaker can receive immediate feedback from the listener and react appropriately. The listener also depends on contextual clues in decoding the speaker's messages. As examples of contextual clues, Georgakopoulou and Goutsos (1997 : 34) cite "visual clues such as body language and gestures, auditory clues such as variations in stress and tone of voice, hesitations, pauses etc.". We can therefore say that spoken discourse is 'contextualised', and written discourse is 'decontextualised'.

These different contextual characteristics have an effect on the linguistic features of both spoken and written discourses. Regarding this matter, Nunan (1995 : 83) states the following:

> the contexts and purposes for using language are inseparable from the language itself. It therefore follows that spoken and written language, which exist to fulfil different functions, will exhibit different characteristics.

Now let us look at some linguistic features of spoken and written language. Citing Richards (1983 : 224-26), Omaggio (1993 : 166-7) notes some of the features which distinguish spoken from written texts as shown below. Although Richards' comments are not satisfactorily evaluative, the basis of his observations is nonetheless accurate.

1. .... Written discourse is normally constructed in sentences, whereas the major constituent in the planning and delivery of spoken discourse is the clause.

2. .... Whereas written discourse typically consists of well-formed sentences, spoken discourse can often include ungrammatical or reduced forms, dropped words, and sentences without subjects, verbs, auxiliaries, and other parts of speech.

3. In well-written discourse, sentences flow in logical sequence and there is evidence of planning of thought. In spoken conversational discourse, pauses, hesitations, false starts, and corrections make up between 30 and 50 percent of what is said. In addition, speakers tend to use fillers and silent pauses to "buy time" as they plan what they want to say next.

4. ... [W]riting tends to be more planned and tightly organized. A written text is usually produced by one person, allowing the discourse to flow logically as the topic is developed. Conversational speech, on the other hand, is generally not planned and therefore not as organized as written discourse. Often there are topic shifts, since the development of the topic of conversation is cooperatively

constructed.

5. Because conversations are interactive, ... [m]any things may be left unsaid because both parties assume some common knowledge. In many types of written discourse, however, ... more background information may be needed in order to communicate clearly.

It is often said, as we have seen above, that spoken language consists of fragmented ungrammatical sentences, false starts, and linguistic redundancy, and that speech is less structured and less highly organised. However, as Halliday (1985 : 76) indicates, although "[t]here is a tradition of regarding spoken language as formless and featureless", this is "a myth" (ibid.: 100). He (ibid.: 79) states that "[t]he spoken language is, in fact, no less structured and highly organised than the written", and "each is complex in its own way. Written language displays one kind of complexity, spoken language another" (ibid.: 62). According to him (ibid.: 87), "[t]he complexity of the written language is static and dense", and "[t]hat of the spoken language is dynamic and intricate".

Before looking at one linguistic feature of differences between spoken and written language, that is, lexical density, it is necessary to note that a clear distinction between them cannot be drawn. As for the distinction between spoken and written language from the point of view of physical conditions, things may now be changing. As stated in Halliday (1985 : 81), the distinction is being made vague, with the development of technology. According to him, in the case of using tape repeaters and transcribing machines, we are able to listen to small chunks of speech repeatedly. While on the other hand, in using computers, we skim and digest our reading matter in chunks by moving the text on the screen. Since in the case of the computer screen, the text is read in a continuous, vertical, top to bottom fashion, the amount of text visible at any one time is limited by the size of the screen. He (ibid.) states that due to modern technology these characteristics of both kinds of discourse change the spoken activities into a kind of "thing" and the writing activities into a "process".

It seems to be also impossible to separate spoken discourse and

written discourse clearly in terms of their respective forms. We can easily detect linguistic characteristics which are often common to spoken discourse and informal conversations in texts sent by electronic mail. On the other hand, in academic lectures, the language used seems to be a written language. In this respect, Nunan (1995 : 84) points out the following:

> Rather than being separate manifestations of language, it has been suggested that spoken and written language exist as a continuum, a mode continuum, and that any given text, spoken or written will exist somewhere along this continuum, depending on the extent to which it exhibits the characteristics of the different forms.

Stubbs (1996 : 74) states that "[t]he clearest difference is not between written and spoken language, which overlap considerably". According to him (ibid.), within the spoken language genres, the absolute difference lies between those in which there is no possibility of feedback from other speakers (e. g. radio commentary, telephone answering machines); and those in which there is feedback (e. g. conversation, radio discussions), and those where there is at least the possibility of feedback (e. g. lectures and speeches) where the audience is physically present.

3. Lexical Density

In this section, I will consider one of the main different linguistic features between spoken and written language: lexical density. The lexical density of the text is defined as the proportion of lexical words, to the total number of words in a text. It can be represented by first, counting the total number of words in the text, and then, after counting the number of lexical words (excluding grammatical words), by expressing the ratio of lexical words to the total number of words as a percentage. "[T]he higher the percentage, the higher the lexical density" (McCarthy, 1990 : 71).

Lexical density can be shown as follows:

lexical density $= 100 \times L/N$.

Where "N is the number of words in a text, and L is the number of lexical words" (Stubbs, 1996 : 72).

As we have seen above, English words can be roughly divided into two categories: lexical words (or items) and grammatical words (or items). Lexical words are also known by the name of 'content (or full) words'. They express content and convey information. In the English language, to express concreteness, nouns, adjectives, adverbs, and main verbs belong to the category of lexical words. Grammatical words are, on the other hand, referred to as 'function (or form or empty) words'. Their function is to relate lexical words to each other. Belonging to the group of grammatical words are auxiliary verbs, modal verbs, pronouns, prepositions, determiners, and conjunctions.

In addition, it is often said that lexical words are an open system, while grammatical words are a closed system. The number of grammatical words is usually limited. For example, according to the *Collins Cobuild English Dictionary* (1995 : 101), "the basic auxiliary verbs are 'be', 'have', and 'do'". New items cannot be added to the group of grammatical words. On the other hand, the number of content words is infinite. We can say that new lexical words can be added and the total number can be expanded infinitely.

Between lexical words and grammatical words, there is no clear dichotomy; there is also a continuum from lexis into grammar. We sometimes find intermediate cases since some kinds of words are midway on the continuum. According to Halliday (1985 : 63), in the English language, prepositions and certain classes of adverb such as 'always' and 'perhaps' are borderline cases. However, as he (ibid.) goes on to say, "[f]or purposes of comparing spoken and written English it does not matter exactly where we draw the line provided we do it consistently".

The difference in the lexical density between spoken and written language is that "[r]elative to each other, written language is dense, spoken language is sparse" (Halliday, 1985 : 62). Ure (1971 : 445) demonstrates that the lexical density of spoken texts is under 40%, while that

of written ones is 40% and over.

However, in Stubbs' research (1996 : 72-3), by examining a larger corpus, he concludes as follows: although, his data for written language agrees with Ure's findings (most written texts are over 40%, the range is 40 to 65%), his spoken language corpus contains a wider range (34 to 58%). A large overlap exists in lexical density between spoken language and written language. As we have seen before, according to Stubbs (ibid.: 74), the clearest difference is between those where there is a possibility of feedback from other participants (varying between 34 and 44%) and those where there is no possibility of feedback (varying between 46 and 64%).

As for the possibility of feedback, Ure (1971 : 448) also claims it is a more important factor in determining lexical density than the choice of the spoken / written language. The dividing line drawn by her is between 36% or under and 37% or over. The texts in which there is a possibility of feedback have a density of 36% or under, while ones in which there is no possibility of feedback have a density of 37% or over.

McCarthy (1990 : 72) points out that the distinction between speaking and writing affects lexical density, however, not all spoken discourse has lower lexical density than that of written discourse: some modes of speech such as oral narratives and formal lectures might be lexically quite dense.

As we have seen above, the lower lexical density of spoken language with feedback can be attributed to the context of production. Since speaking is a faster activity in which the participants do not always have enough time to organise and elaborate their utterances, spoken language includes repetition. Moreover, contextualised characteristics of spoken discourse cause low lexical density. According to the table in Ure (1971 : 451), language in action (e. g., immediate doing, consulting, non-immediate describing, directions, and discussion) has lower lexical density. In language in action, "[t]here is ... no need for elaboration: a lot of knowledge is shared, and a lot of referents can be mutually taken for granted" (McCarthy and Carter, 1995 : 209). By taking the above facts into consideration, we can expect the lexical density in casual con-

versations to be the lowest.

4. Methodology

In this section, I will analyse some dialogues appearing in EFL textbooks used in Japanese senior high schools in terms of lexical density. The aim of the analysis is to examine whether the texts used in the EFL textbooks aimed at promoting speaking ability have the characteristics of spoken discourse discussed in the earlier section. I will therefore focus my attention on lexical density.

4.1 Choice of Textbooks

The textbooks used in this study are all government authorised textbooks used in Japan. They are used for "oral communication A", classes which aim to develop speaking ability, in senior high schools. There are sixteen textbooks, and details of these are given in Appendix 1.

4.2 Selecting Texts

For the analysis of lexical density, I have selected, as data, texts which are fully transcribed in the pages of the students' textbooks: this means I have disregarded texts, for example, used in listening comprehension exercises because they omit vocabulary as part of the question format. The principal focus of this kind of exercise may not be on the overall structure of the dialogue as a model, but on the words or expressions which are omitted.

In addition, even if the text is fully transcribed, texts which consist of three utterances or less are also ignored, since they are too short to provide accurate assessment, and may be also used as a vehicle for teaching specific words, phrases, or expressions. Several textbooks such as *Select* and *Expressways* have sections for teaching various expressions in the introductory section. For example, in *Expressways,* a situation in which conversation is occurring is explained first and then a short exchange between two persons is given:

LESSON 7 SUGGESTING

Sarah is visiting Ken's house. They are discussing what to do on Sunday.

Sarah: What shall we do today?

Ken: How about going to Disneyland?

Sarah: Okay. Good idea! (p. 14)

Following this conversation, four exercises are given: first, pupils listen to the questions about this conversation and answer the questions; then, they expand the basic dialogue by taking turns and adding three sentences each. Then, in the next exercise the pupils complete each set of sentences by using the given words, in this lesson, the target expression is 'How about...?', and set phrases such as 'going to the movies' and 'going shopping in town' are given; finally, they have to match the phrases with the responses. We can see that the main purpose of this lesson is to teach the pupils how to form expressions using the 'How about...?' construction.

In four textbooks (*Departure, Evergreen, Interact,* and *Speak to the World*) out of the sixteen, we cannot find any fully transcribed texts, while in twelve textbooks, fully transcribed dialogues are used as models. Texts appearing in these textbooks are all informal conversations between two or three participants, either face-to-face conversations or telephone conversations. Although each textbook has its own format and content, some topics and situations set for conversation are common to most of the textbooks. The conversations dealing with the linguistic functions, used in 'giving / asking directions', and 'talking about likes and dislikes' seem to be common to many of the textbooks, therefore I will focus my attention on the texts which use these situations.

As for the conversations dealing with the linguistic functions used in 'giving / asking' directions, I will look at thirteen texts as shown in Table 1 below. In eleven textbooks out of the twelve which have fully transcribed texts, we find dialogues containing constructions used in 'giving / asking directions'. One textbook does not have a dialogue containing such constructions, while two textbooks (*Select* and *The Crown*) each have two such dialogues. As for the conversations dealing with the

linguistic functions used in 'talking about likes and dislikes', there are twelve texts from nine textbooks, which are shown in Table 2.

4.3 Counting Lexical Words

Concerning the distinction between lexical words and grammatical words, I showed in the preceding section that much discussion on this unclear dichotomy exists and that borderline cases can be found in the English language. In this study, I will use the following basic distinction:

> Lexical words: nouns, (open-class) adjectives, (open-class) adverbs, and main verbs.
> Grammatical words: auxiliary verbs, modal verbs, pronouns, prepositions, determiners, and conjunctions.

Also, as for the grammatical (or function) words, I will follow the definition provided in *The Kenkyusha Dictionary of English Linguistics and Philology* (1987 : 462-3). Most of the definitions given below, of course, overlap with the definition given above.

> A: Determiners: articles, possessive pronouns, quantifiers, demonstratives, possessive proper nouns. 'Each', 'both', etc. can be included in this group. They are used in the same position as 'the' in '*The* concert was good'.
> B: (a) 'may', 'might', 'can', 'could', 'will', 'would', 'should', 'must', 'has (been)', 'has to (be)', used in the same position as 'may' in 'The concert *may* be good'.
> (b) 'had', 'was', and 'got', used in the same position as 'had' in 'The student *had* moved'.
> (c) 'was', 'got', and 'kept', used in the same position as 'was' in 'The student *was* moving'.
> (d) 'had to' and 'did', used in the same position as 'had to' in 'The student *had to* move'.
> C: 'not', as in 'The concert may *not* be good'.
> D: (a) 'very', 'quite', 'awfully', 'really', 'awful', 'real', 'any', 'pretty',

'too', 'fairly', 'more', 'rather', and 'most', used in the same position as 'very' in 'The concert may not be *very* good then'. Most of them are used as an intensifier or a down-toner.

(b) 'still', 'even', 'much', 'some', and 'no', used in the same position as 'still' in 'The student may be *still* better'.

E: 'and', 'or', 'not', 'nor', 'but', and 'rather than', used in the same position as 'and' in 'The concerts *and* the lectures are *and* were interesting *and* profitable now *and* earlier'.

F: Words which can appear in the same position as 'at' in the sentence 'The concerts *at* the school are *at* the top'.

G: 'do', used in the same position as 'do' in '*Do* the boys do their work promptly? / The boys *do* not do their work promptly'.

H: 'there', as in '*There* is a man there. / Is *there* a man there?'

I: 'when', 'why', 'where', 'how', 'who', 'which', and 'what', used in the same position as 'when' in '*When* was the concert good?'

J: 'after', 'when', 'so', 'whenever', 'because', 'therefore', 'although', 'and', 'since', 'but', and 'before', used in the same position as 'after' in 'The orchestra was good *after* the new conductor came'.

K: 'well', 'oh', 'now', and 'why': which appear at the beginning of the response utterances.

L: Words which appear at the beginning of the response utterances, such as 'yes' and 'no'.

M: Words used in the situation utterances to initiate the conversation, such as 'say', 'listen', and 'look'.

N: 'please', as in '*Please* take these two letters'.

O: 'lets (=let's)', as in '*Let's* do the invitations right away'. (summarised, translated, and emphasis added)

Other rules used in counting words in the texts are as follows: (1) a proper noun consisting of two words or more is counted as one word, therefore, "Sunrise Hotel" for example, is counted as one noun; (2) contractions are counted as one word; (3) hyphenated words are counted as one, and (4) utterances in Japanese are excluded from the total number of the words in the text, for example, "*Eigo Wakarimasen*. (I don't under-

stand English.)".

Here is a sample dialogue which shows the principles in counting lexical / grammatical words. Words in bold indicate lexical words, underlined words are counted as one word, and words in italics are excluded from the total number of words in the text.

Text in *Select* (p. 100)

Judy: **Excuse** me, where is <u>**Nishi Station**</u>?

Stranger 1: *Eigo wakarimasen.*

Judy: **OK, thank** you.

Stranger 2: Can I **help** you?

Judy: **Thank God!** Yes, please! Will you **tell** me the **way** to <u>**Nishi Station**</u>?

Stranger 2: **Go straight** and **turn left** at the **corner.** <u>It's</u> **next** to the **big Department Store.**

Judy: **Thank** you very much. You **saved** me.

Stranger 2: <u>You're</u> **welcome. Have** a **nice day**! Bye!

Judy: Bye!

Total number of words: 54

Lexical words: 25

Lexical density: 46.3% (rounded off to one decimal place)

4.4 Results and Discussion

4.4.1 Giving / Asking Directions

Table 1 shows the lexical density of the texts, the situation being 'giving / asking directions'. By looking at Table 1, we can see clearly that the lexical density of the texts appearing in the EFL textbooks is fairly high, except the ones in the text found in *Select* (31.6%), in *Progressive* (37.2%), and in *Mainstream* (43.8%). As we have seen in the above section, by examining a large corpus, Stubbs' research (1996) shows that the lexical density of the texts in which there is a possibility of feedback from other participants varies between 34 and 44%. In Ure (1971), it is also stated that the texts in which there is a possibility of

Lexical Density of Invented Dialogues in EFL Textbooks

Table 1 Giving/Asking Directions

Title of textbook	Words	Lexical words	Lexical density
Select (p. 100-a)	38	12	31.6%
Progressive (p. 54)	43	16	37.2%
Mainstream (p. 62)	72	32	43.8%
English Street (p. 75)	101	45	44.6%
On Air (p. 13)	66	30	45.5%
Select (p. 100-b)	54	25	46.3%
The Crown (p. 24)	51	24	47.0%
Birdland (p. 28)	90	44	48.9%
The Crown (p. 18)	50	25	50.0%
New Start (p. 48)	68	34	50.0%
Expressways (p. 64)	52	26	50.0%
Echo (p. 28)	61	32	52.5%
Hello, there (p. 74)	87	48	55.1%

Percentages are rounded off to one decimal place.

feedback have a density of 36% or less. All the texts except three have higher density than that mentioned in these studies, despite the fact that there is a possibility of feedback from other participant(s): in fact, these texts revolve around a conversation between two people.

Now, let us look at the text which has the lowest density (text (1)) and the text which has the highest density (text (2)), in order to compare the difference between the two. The sentence patterns asking about the way and expressing gratitude are very similar in both texts. The difference may lie in the body of each conversation. Although text (1) is very short compared with text (2), it seems that the use of the pronoun 'it' in text (1) lowers the lexical density. In the case of genuine conversation, McCarthy (1990 : 72) states, that as "[t]he two speakers... had all the physical context in front of them, ... a higher number of ... pronouns are used". From this point of view, sentences used in text (1) seem to be more natural: about one quarter of the words in the text are pronouns. Appendix 2 and 3 show all the lexical and grammatical words in each category with respect to each dialogue as given below.

(1) Text in *Select* (p. 100-a) Lexical density: 31.6%

Dan has lost his way and is looking for the hotel where he is staying. (translated)

Dan: Excuse me, will you tell me the way to the Sunrise Hotel?

Policeman: It's at the end of the block. You'll find it on your right. Look, you can see it from here.

Dan: Thank you very much.

Policeman: You are welcome.

(2) Text in *Hello, there* (p. 74) Lexical density: 55.1%

A foreign woman is asking Kumi directions in front of a shop on Harajuku Street. (translated)

W: Excuse me. Can you tell me the way to Takeshita Dori Street?

K: Sure. Go straight along this street until you come to a big street. That's Omotesando Street. Then turn left.

W: All right. Go straight, and then turn left at the big street.

K: Keep going until you come to a big intersection and turn right.

W: Then turn right at the big intersection.

K: O. K. Then keep going until you come to a smaller intersection. Cross the street and turn left. You'll be on Takeshita Dori Street. You can't miss it.

W: Thanks a lot.

K: You're welcome.

One problem with much of the data used here is the shortness of the text itself. For example, because text (1) itself is few in words, changing just one word can have a large effect on the lexical density. Here, it is worth looking at the text in *English Street* (p. 75), which has the most words and comparatively lower density.

(3) Text in *English Street* (p. 75) Lexical density: 44.6%

A: Excuse me...

B: Yes?

A: How can I get to North High School?

B: OK. Go up this street to the first light. That's Central Avenue. Turn left and keep going until you come to the river.

A: So I turn left at Central and go as far as the river?

B: That's right. Then turn right on East River Road and continue until you come to the 2nd Avenue Bridge.

A: The 2nd Avenue Bridge?

B: Yes. Cross the bridge and you'll see North High across the street on your right. You can't miss it.

A: Is it too far to walk?

B: Well ... it's about a 20-minute walk.

A: That's not so far. Thank you very much.

B: You're welcome.

Although there seem to be no differences in the structure of sentences and lexical items used both in text (2) and (3), text (3) sounds more natural. On comparison of text (3) and text (2), the difference may lie in the use of adverbs: in text (3), we can find ten adverbs, whereas only a few are used in text (2). Text (2) sounds artificially constrained and more densely packed, because it lacks markers such as 'well' and 'so' that make a dialogue sound more natural. By adding some function words such as markers and fillers, the transcribed conversation sounds more natural without changing the structure of the target sentences and the target words or expressions.

It may also be worth noticing that text (2) is more informative, and therefore it has more repetition of lexical items—this makes it more lexically dense. In text (2), participants repeat the same words in the same order, while in text (1) and text (3), pronouns, paraphrases, and ellipses are used in response utterances. This causes higher density in text (2), and may lead to a feeling of unnaturalness, despite the informal nature of the face-to-face conversation.

4.4.2 Talking about Likes and Dislikes

In this part, I will examine another type of conversation: talking about likes and dislikes. Table 2 shows the lexical density of the texts in

Table 2   Talking about Likes and Dislikes

Title of textbook	Words	Lexical words	Lexical density
Birdland (p. 21)	54	23	42.6%
Progressive (p. 10)	36	16	44.4%
Hello, there (p. 34)	51	23	45.1%
On Air (p. 31)	53	24	45.3%
Mainstream (p. 36)	78	37	47.4%
Progressive (p. 32)	84	40	47.6%
On Air (p. 17)	77	38	49.4%
Mainstream (p. 34)	68	34	50.0%
Hello, there (p. 42)	66	34	51.5%
Select (p. 98)	69	36	52.2%
Echo (p. 12)	56	30	53.6%
English Street (p. 79)	122	70	57.4%

Percentages are rounded off to one decimal place.

which topics about likes and dislikes are dealt with.

We also find here that the lexical density of each text is quite high. One of the reasons for this may be the omission of the opening and closing part of the conversation. Text (4) has a density of 41.8%, in spite of the use of function words such as 'well', 'oh', and 'Mmm'. Almost all the texts used in this study start with the questions 'Do you like...?', 'What... do you like', and so on. Actual everyday conversations rarely start with just asking questions, however. In a series of conversations, we can find structures consisting of opening, main body, and closing; and conversations are composed of various items apart from simple sentence statements. Most of the texts analysed here seem to be intended simply to teach expressions for talking about likes and dislikes.

(4) Text in *Birdland* (p. 21) Lexical density: 42.6%

Ms. Sato: What do you think about this symphony? It's beautiful, isn't it?

Ken: Well, I think it's a little boring.

Ms. Sato: Oh, why do you think so?

Ken: I don't think it has any punch to it.

Ms. Sato: Mmm... What kind of music do you like?

Ken: I like rock music.

Ms. Sato: I see. I hope someday you'll like classical music, too.

(5) Text in _English Street_ (p. 79) Lexical density: 57.4%

Bob: What do you like to do in your free time?

Kyoko: I like listening to music, pop songs. I like foreign singers.

B: Foreign singers?

K: Yes. I like American pop singers. My favorite is Mariah Carey. She's great. Do you listen to pop songs much?

B: Sometimes... but I really prefer jumping and running to listening to music.

K: You mean sports?

B: Yes, I like playing tennis, swimming, going on hikes... and playing cards, too.

K: Playing cards? That doesn't fit your "athletic image".

B: Oh, well, sometimes I like to waste time. What do you like besides listening to music?

K: I also like going cycling. I often go to the beach or to the mountains...

B: I like cycling, too. Maybe we could go together sometime.

K: I'd like that.

It is noteworthy that in text (5), which has the highest lexical density, a listing of lexical items often appears: for example, 'jumping and running' and 'playing tennis, swimming, going on hikes... and playing cards', whereas in text (4) there are no lists of vocabulary in the same category. This feature of text (5) is one of the reasons for having higher lexical density.

In EFL textbooks, lists of vocabulary items in the same categories are often used, as in text (5). There seems to be the intention to give learners as many vocabulary items as possible in the belief that this supplements learning. However, as stated above, this brings about high lexical density, and this may be one of the reasons for the unnaturalness of the dialogue as a model of face-to-face informal conversations.

In addition to the lists of lexical items, we find that certain sentences are repeatedly used to express 'likes': for example, 'I like listening to music', 'I like foreign singers', and 'I like to waste time' in text (5). In fact, text (5) comprises the repetition of the question-answer format regarding their favourite things. This text sounds very informative and this may lead to higher lexical density.

By examining the lexical density of the texts used in EFL textbooks, we have found that (1) the lexical density of dialogues in EFL textbooks is relatively higher than that of genuine everyday face-to-face conversations; (2) in the EFL textbooks studied here, the use of pronouns and adverbs seems to be responsible for the difference in lexical density. (3) One of the reasons for higher density is that opening and closing parts are often omitted because of the pedagogic intention of teaching useful words and expressions. Also, (4) other reasons for higher lexical density are the listing of lexical items and the repetition of the question-answer format which contains the target expressions.

5. Implications for Language Teaching

In the previous section, we have seen that differences exist between spoken and written language, and that they have their own characteristics. Although in language education, written language seems to have been taught, as Stubbs (1996 : 64) states, "written language is not simply spoken language written down", and equally, spoken language is not simply written language read aloud.

As a result of the criticism that students do not have the opportunity to use English in practical situations, especially in terms of spoken English, nowadays more attempts to focus attention on teaching speaking in English are being made in Japan, including the introduction of subjects such as 'oral communication' into the curriculum of senior high schools. However, the texts used in the classroom aimed at improving students' speaking ability do not incorporate the linguistic characteristics which are peculiar to spoken language. From the analysis of the text-book dialogues in terms of lexical density, which is one of the differences between spoken (face-to-face informal conversational) and written lan-

guage, it was found that the lexical density of the invented dialogues in the textbooks is as high as that of written language.

In several EFL situations, like in Japan, most learners receive target language input only in the language learning classroom environment. In these circumstances, materials used in classes are essential, and above all, textbook dialogues play a crucial role. Almost every textbook contains dialogue(s) in each lesson or part, and these dialogues are often presumed to be models of real-life English. There are certain structured activities using these dialogues; for example, in team teaching, the Japanese English teacher reads the dialogue with an assistant English teacher; students read the dialogue working in pairs; or students act out the dialogue in front of the class from memory. Although other kinds of input may be provided in the class, for learners, especially in the early stages, the textbook dialogue is the main vehicle for learning to communicate in English. From this viewpoint, we can say that textbook dialogues used in spoken English classes, should reflect the features of spoken language.

One of the major characteristics of the invented dialogues with higher lexical density in EFL textbooks is the lack of function words such as fillers, 'Mmm', 'Well', and 'Erm'. Although it is questionable whether these words should be taught as lexical items with clear instructions in the classroom, in the texts which are used as models of conversations, these characteristics should be presented in order to make the learners aware of the strategies used by native speakers of English to gain time to think and plan what they are going to say.

As has been pointed out, incompleteness and ungrammaticality of sentences are two of the characteristics of spoken language. However, in the teaching environment, this does not mean that it is necessary for the learners to be taught to speak in incomplete and ungrammatical sentences. It might be that the features and the structure of spoken discourse, as we have seen above, should be taught, or at least presented in textbooks, so that, at least, teachers are made aware of them.

6. Conclusion

In this paper, invented dialogues appearing in EFL textbooks have been analysed on the basis of the findings apparent in the study of discourse, in terms of lexical density. Although many features of naturally occurring conversations have been demonstrated in investigations during the past few decades, these findings do not seem to be reflected in EFL textbooks which aim to foster the speaking ability of pupils in senior high schools in Japan.

I analysed invented dialogues in EFL textbooks in terms of their lexical density, on the assumption that the lexical density of naturally occurring conversations in which there is a possibility of feedback by the other interlocutor(s) is lower (to be exact 44% or under) than that of other kinds of texts where there is no possibility of feedback, including, of course, written discourse (Ure, 1971 and Stubbs, 1996). Results indicate that the lexical density of the texts appearing in the EFL textbooks is considerably high, despite the fact that dialogues analysed in this study are texts consisting of face-to-face conversation, topics in which 'giving / asking directions' and 'talking about likes and dislikes' feature predominantly.

By comparing the text which has the lowest lexical density and that which has higher lexical density, it was found that texts that have relatively high lexical density (1) lack the use of pronouns, (2) lack the use of markers such as 'well' and 'so,' and (3) are more informative, therefore lexical items are often repeated. In addition to this, most of the textbook dialogues omit the opening and closing part of the conversation, and this makes the texts more lexically dense.

In certain EFL situations, like in Japan, the choice of materials used in the classroom is essential, particularly textbook dialogues. These dialogues are often used as models of real-life English, and other practice activities are often based on them. Although other kinds of input may be provided in the class, for beginners, textbook dialogues are of prime concern. Clear detailed instructions regarding the features of spoken language or the structure of the conversation may not be needed in the

classroom, however, they should, as we have seen in this paper, be at least presented in the textbook in order to make teachers aware of these features.

Notes

1. An earlier version of this paper was submitted to the University of Birmingham as part of my M.A. dissertation. I would like to express my gratitude to Dr. Susan Hunston for her comments and suggestions.
2. From 2003, new guidelines will take effect, and in these new study guides, a statement reflecting the new aims and concerns of the Japanese Ministry of Education will include the intention of, "developing pupils' practical ability to communicate in English" (Monbusho, 1999 : 119).

References

Biber, D., S. Conrad, and R. Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Georgakopoulou, A. and D. Goutsos. 1997. *Discourse Analysis: An Introduction*. Edinburgh: Edinburgh University Press.

Halliday, M. A. K. 1985. *Spoken and Written Language*. Oxford: Oxford University Press.

McCarthy, M. 1990. *Vocabulary*. Oxford: Oxford University Press.

McCarthy, M. 1991. *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.

McCarthy, M. and R. Carter. 1995. 'Spoken Grammar: What Is It and How Can We Teach It?'. *ELT Journal* 49/3 : 207-218.

Monbusho. 1989. *The Course of Study for Upper Secondary Schools*. Tokyo: Monbusho.

Monbusho. 1999. *The Course of Study for Upper Secondary Schools*. Tokyo: Monbusho.

Nunan, D. 1993. *Introducing Discourse Analysis*. London: Penguin Books.

Nunan, D. 1995. *Language Teaching Methodology: A Textbook for Teachers*. Phoenix ELT.

Omaggio, A. 1993. *Teaching Language in Context*. Heinle & Heinle.

Otsuka, T. and F. Nakajima (Eds.) 1987. *The Kenkyusha Dictionary of English Linguistics and Philology*. Tokyo: Kenkyusha.

Perren, G. and J. L. M. Trim (Eds.) 1971. *Applications of Linguistics*. London: Cambridge University Press.

Richards, J. C. 1983. 'Listening Comprehension: Approach, Design, Procedure'. *TESOL Quarterly* 17 : 219-40.

Sinclair, J. (Ed.). 1995. *Collins Cobuild English Dictionary,* 2nd. edn. London: HarperCollins.

Stubbs, M. 1996. *Text and Corpus Analysis.* Oxford: Blackwell.

Tribble, C. 1996. *Writing.* Oxford: Oxford University Press.

Ure, J. 1971. 'Lexical Density and Register Differentiation'. In G. Perren and J. L. M. Trim, (Eds.): 443-52.

APPENDIX 1   List of Textbooks Used in the Study.

Birdland: Oral Communication A. 1997. Bun-eido.

The Crown: Oral Communication A. 1998. Sanseido.

Departure: Oral Communication A. 1998. Taishukan.

Echo English Course: Oral Communication A. 1998. Sansyusha.

English Street: Oral Communication A. 1998. Daiichi Gakushusha.

Evergreen: Communication A. 1998. Daiichi Gakushusha.

Expressways: Oral Communication A. 1997. Kairyudo.

Hello, there! : Oral Communication A. 1997. Tokyo: Tokyo Shoseki.

Interact: Oral Communication Course A. 1998. Kirihara Shoten.

Mainstream: Oral Communication A. 1998. Zoshindo.

New Start English: Communication A. 1997. Obunsha.

On Air: Communication A. 1998. Kaitakusha.

Progressive: Oral Communication A. 1998. Shogaku Tosho.

Select: Oral Communication A. 1998. Sanseido.

Speak to the World: Oral Communication A. 1994. Kyoiku Shuppan.

Talk up: Oral Communication A. 1997. Keirinkan.

APPENDIX 2   List of Lexical/Grammatical Words Found in Texts(2-1)-(2-3)

	(2-1) Select (100-a)	(2-2) Hello, there ! (74)	(2-3) English Street (75)
Grammatical words			
Pronouns	me [2] you [4] + you ('ll) it [2] + it ('s)	me [2] you [5] + you ('re), you ('ll) it	me I [2] you [4] + you ('ll), you ('re) it [2] + it ('s)
A Determiners	your the [4]	the [4] a [3] that ('s) this	your the [7] a that ('s) [3] this
B Modal and auxiliary verbs	will can are	can + can ('t) be keep [2]	can + can ('t) is keep
D Adverbs	very much	a lot	very much how up too not so
E/J Conjunctions		until [3] and [3]	and [4] until [2] so
F Prepositions	to at of on from	to [4] at [2] on along	at on [2] to [5] across as
K			well
L			yes [2]
M	look		
TOTAL	26	39	56

Lexical words			
Nouns	way Sunrise Hotel end block here	way Takeshita Dori St.[2] street [4] Omotesando St. intersection [3]	North High School North High street [2] light Central Avenue Central river [2] East River Road 2nd Avenue Bridge[2] bridge walk
Adjectives	right welcome	big [4] smaller welcome	first right [2] 20-minute welcome
Adverbs		sure straight [2] then [4] left [3] all right right [2] O. K.	OK left [2] as far [3] then right about
Main verbs	excuse tell find see thank	excuse tell go [2] + going [2] come [3] turn [5] cross miss thanks	excuse get go [2] + going turn [3] come [2] continue cross see miss walk thank
TOTAL	38	48	45

Alphabetical letters in bold correspond to the definitions given in 4. 3.
Numbers in square brackets correspond to frequency of use in the text.

Lexical Density of Invented Dialogues in EFL Textbooks

APPENDIX 3 List of Lexical/Grammatical Words Found in Texts(2-4)-(2-5)

	(2-4) Birdland (21)	(2-5) English Street (79)
Grammatical words		
Pronouns	you [3] + you ('ll) it [3] it ('s) [2] I [5]	you [4] I [9] + I ('d) she ('s) we
A Determiners	any this	your [2] my the [2] that [2]
B Modal and auxiliary verbs	is (n't)	is could
D Adverbs	a little	much
E/J Conjunctions		but and [2] or
F Prepositions	to of about	to [9] in on besides
G	do [3] + do (n't)	do [3] does (n't)
I	what [2] why	what [2]
K	well oh mmm	oh well
L		yes [2]
TOTAL	31	52
Lexical words		
Nouns	symphony punch kind music [3] rock	time music [3] songs [2] singers [3] favorite

	classical	—	Mariah Carey sports tennis hikes cards [2] image time cycling [2] beach mountains
Adjectives	beautiful boring		free pop [3] foreign [2] American great athletic
Adverbs	so someday too		sometimes [2] really too [2] also often maybe together sometime
Main verbs	think [4] has like [3] see hope		like [10] do listen + listen (ing) [3] prefer jump (ing) run (ning) mean play (ing) [3] swim (ming) go [2] + go (ing) [2] fit waste
TOTAL	23		70

Alphabetical letters in bold correspond to the definitions given in 4. 3.
Numbers in square brackets correspond to frequency of use in the text.

Lexical Density of Invented Dialogues
in EFL Textbooks:
How Natural Are They
as Models of Spoken Discourse?

TSURII, Chie

This paper attempts to investigate how natural invented dialogues appearing in EFL textbooks are as models of informal conversations. I analyse texts in EFL textbooks aimed at improving the speaking ability of students in senior high schools in Japan, in order to investigate the differences between invented dialogues and authentic conversations. The texts are analysed and discussed in terms of their lexical density, on the assumption that the lexical density of naturally occurring conversations in which there is a possibility of feedback by the other interlocutor(s) is lower than that of other kinds of texts where there is no possibility of feedback, including written discourse.

By examining the data collected from twelve textbooks, I have found that the lexical density of the texts appearing in a selection of textbooks is considerably high, despite the fact that dialogues analysed in this study are texts consisting of face-to-face conversations in which there is a possibility of feedback by the other interlocutor(s). As for the reasons behind the higher density of invented dialogues, these appear to be: (1) the use of fewer pronouns and adverbs; (2) the lack of opening and closing parts; and (3) the listing of the lexical items and the repetition of the question-answer format which contains the target expression. In several EFL situations, like in Japan, materials used in the classroom play a crucial role in learning performance, therefore, the discourse features of colloquial English should be at least presented in textbooks. Teachers should also be made aware of these factors.