東北公益文科大学
総合研究論集

17

An analysis of student use of formulaic sequences
in an EFL conversation test

Adrian Paterson

東北公益文科大学
Tohoku University of Community Service and Science

研究論文

An analysis of student use of formulaic sequences in an EFL conversation test

Adrian Paterson

This paper describes the process of developing and implementing a test of EFL students' English conversation ability with a particular focus on formulaic sequences. Formulaic sequences are common phrases that are stored and retreived in much the same ways as individual words. It, therefore, follows that they could also be taught and tested in much the same way as individual words. This and other linguistic theories and research influenced my decisions regarding textbook selection and test development, and these will be discussed along with some pedagogical issues related to teaching vocabulary and formulaic sequences in an EFL class. The results of the EFL conversation test show that there is considerable variation in students' prior knowledge of formulaic sequences and conversation strategies, and the ease or difficulty with which they acquire them.

Formulaic sequences

Formulaic sequences are groups of 2 or more words that behave like an individual word, such as; commonly used expressions, phrasal verbs, noun phrases, idioms, etc. They are not always included in definitions of vocabulary. However it is important that they be considered part of the lexicon of a language, because there are many which have to be learned as a whole and cannot be understood simply by looking at the individual words from which they are made up. Read (2000) gives some examples; "phrasal verbs (*get across, move out, put up with*)", "compound nouns (*fire fighter, love letters, practical joke, personal computer, applied social science, milk of magnesia*)", and "idioms (*a piece of cake, the good book, to go the whole hog, let the cat out of the bag*)" (pp. 20-21). Some further examples are; *How's it going?, to go out* [date]*, Japan Association of Language Teachers* [JALT], *You can lead a horse to water, but you cannot make him drink*.

While linguists have been aware of formulaic sequences for a long time, they were marginalized due the dominance of Chomskian linguistics with its emphasis on generative, rule-based grammatical sentence construction. They have recently undergone a resurgence as an area of interest in vocabulary research, two influential recent works are Wray (2002) and Schmitt (2004). Wray's working definition of formulaic sequences, also adoped by Schmitt, is:

> a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (Wray, 2002, p. 9)

Formulaic sequences allow learners to make use of ready made "chunks", which can help them to take an active part in conversations, and can be learned quickly using rote learning techniques. Rote learning has fallen out of favour with the advent of communicative language teaching, but vocabulary is actually one area of language acquisition where rote learning can be beneficial for developing initial knowledge of a word's form and meaning. Communicative activities then become more effective at developing and expanding knowledge of word usage. It is effective and accurate use of language, in particular formulaic sequences, that is one of the hallmarks of native-like or near native-like language use. Recent corpus-based research into native-speaker use of formulaic sequences is revealing that they make up a significant proportion the language used. Especially, in the case of spoken language.

Conversation Strategies

Kenny and Woo (2004) advocate use of communication strategies, which they define as follows:

> We define conversation strategies broadly as the lexical phrases learners use to maintain their conversations, solve communication problems, and give feedback to their partners. We further define them by describing what they are not: they are not vocabulary items or idioms, they are not topic questions,

indeed, they are not even topic-specific. They are relatively limited in number and mastery of them goes a long, long way. (Kenny and Woo, 2004, p. 4).

The problem with this definition is that it confuses two different things; the speaker's intention, or more specifically the pragmatic function it performs, and the actual means of achieving it. A strategy is a plan of action designed to achieve a long-term or overall goal. In Kenny and Woo's definition the goals are; maintaining conversations, solving communication problems, and giving feedback. In contrast, formulaic sequences, or lexical phrases as they call them, are the means that allow a person to communicate that aim. In most cases, for any particular communicative strategy, there are several formulaic sequences available to the speaker that will work equally well in any given circumstance. Therefore, I think that from a theoretical point of view, it is important to make a distinction between these two different, but closely related concepts. For the purposes of this paper, the word strategy will only refer to the speaker's intention or communicative goals, and formulaic sequences will refer to the actual phrases used to achieve them.

Why do we need formulaic sequences?

The reason speakers need formulaic sequences is that language is very complicated, and real time processing places a heavy burden on the speaker. This can best be illustrated using Levelt's model of speech processing (Levelt, 1989) which offers some useful insights into how we store, process, and produce language (see Figure 1).

Let's imagine that I want to tell someone about a friend called John who gave a present to his girlfriend called Mary, and let's assume that the listener knows both of them. Firstly, I start with a concept in a part of my mind Levelt calls the conceptualizer. This concept could be either a memory or an image of the event depending on whether I saw it or heard about it. To this concept the conceptualizer adds various knowledge about the listener, John and Mary, their relationships, and other knowledge about the world in general. All of this happens to the concept, or pre-verbal message, before it has been put into words. This pre-verbal message then goes to the formulator. The
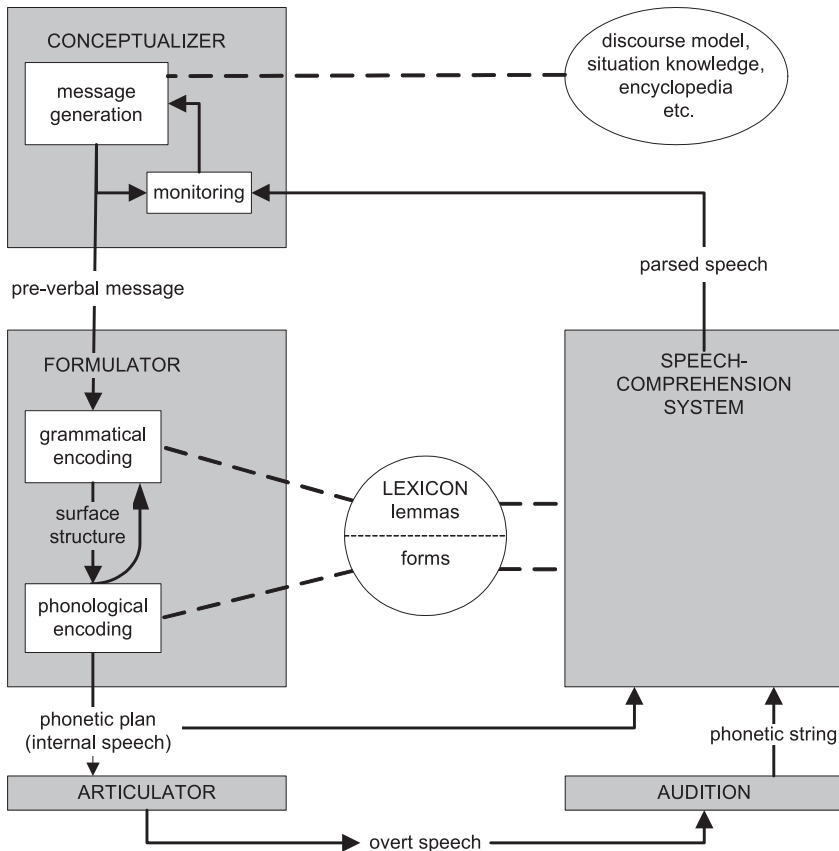
Figure 1. Levelt's Model of speech processing

formulator's job is to: 1) Select words including lemma information from the lexicon or word store (John, Mary, a present, [subject, person] give[third-person, singular, past] [direct object, thing] to [indirect object, person]). 2) Use grammar and collocation rules to arrange the words into a phrase, clause, or sentence (John gave a present to Mary). 3) Use phonological rules to encode the pronunciation (/jongevaprezentomeri/). 4) And then to create a phonetic plan of how to say it. This phonetic plan then goes to the articulator which controls the muscles that produce speech and it becomes overt speech. Notice that monitoring or self correction occur at each step along the way,

either within the conceptualizer or through the speech comprehension system. In the case of fluent speech this is all occurring extreemely fast.

Collocation

In her definition of formulaic sequences quoted above, Wray uses the terms 'continuous or discontinuous' to describe one of their characteristics. What she means is that they can have open slots, into which other words or formulaic sequences can be inserted. In many cases, there are a limited range of words, formulaic sequences, or categories thereof that can be inserted. This is a form of collocation.

Collocations can be very useful for learners. They are words which commonly occur together. They are usually a pair, but are not always next to each other. Some words are more limited in terms of the range of words with which they can collocate than others. For example, in the collocation 'take a bath', 'take' can collocate with many words (e.g. 'Take a cookie', 'Take my son to the doctor', 'Take a picture', 'Take a day off', 'Take a magazine from the shop/ [steal]', 'Take a magazine from the shelf [to look at or buy]', etc.). Whereas, 'bath' is more limited (e.g. 'Clean the bath', 'Fill/ empty the bath', and 'Give someone a bath').

Collocations are sometimes confused with grammar items, particularly in tests. Many so called 'grammar' questions in English tests are actually collocation questions. To illustrate this, let's look at two different sets of multi-choice answers to an exam question "I _____ a bath last night";

took, sold, read, called (collocation)

take, took, takes, taken (grammar)

In the first set of answers, the test taker has to choose the appropriate word to go with 'bath', therefore, it is a collocation question. However, in the second set of answers they have to choose the past tense of the verb to match the adverbial of time 'last night', therefore, it is a grammar question. This is a common problem, and it is very important for test writers to be aware of exactly what they are testing.

Formulaic sequences and collocation are important to second language learners because they give them sets of readymade combinations of language that they can

use with minimal processing thereby helping them to improve fluency. In fact it is the ability to use formulaic language and collocation which is one of the defining features of native-speaker or native-like language. This makes learning them a high priority for learners.

Word-based versus grammar-based language

The importance of formulaic sequences in native-speaker language lends support to Lewis' (1993) argument that a lot of grammar is actually word driven. He argues that "Language consists of grammaticalised lexis, not lexicalised grammar" (1993, vi). In other words, instead of starting from a grammatical pattern and then inserting words, as many textbook writers and publishers would apparently have us believe, we start with a concept, as we saw in Levelt's model, which is then labelled using words that are then arranged in a pattern to convey the intended meaning. So the above example sentence 'John gave Mary a present.' according to Levelt's model, goes through the following process:

Image (concept) → **labels (words)** → **organise (grammar)** → **output (speech/writing)**
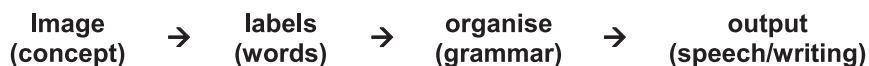
I think that this seems to be a more logical order than one that tries to organise the phrase before adding the words. As we saw in the above example the sentence grammar is determined by the choice of the verb 'give', suggesting that grammar is word driven rather than the other way around. This is not to say that grammar should not be taught, but rather that vocabulary needs to be given a more prominent place in the language classroom. However, a lot of teaching materials tend to overemphasise grammar at the expense of vocabulary.

A further argument for emphasising vocabulary is that many ungrammatical sentences can still be understood. For example, a learner who said 'Yesterday shop go' should be understood by most people to be trying to say 'Yesterday, I went to the shop'. The use of the adverbial 'yesterday' defining the sentence as past, and the verb 'go' and the object 'shop' carrying the key ideas (despite the object being in the subject

position), would allow most listeners to deduce the intended meaning. If such a learner had learned it as a formulaic sequence; 'I went to the ___', they could have produced it more acurately. As David Wilkins so aptly points out "Without grammar little can be conveyed; without vocabulary nothing can be conveyed" (quoted in Lewis, 1993).

Vocabulary acquisition

Because formulaic sequences are more like vocabulary items than grammatical items, it also follows that they can be taught in much the same way. The problem is that it is more difficult to identify high frequency formulaic sequences because it is difficult to program corpus analysis software to find them. This is made particularly difficult by the relatively vague nature of the definition of formulaic sequences. Vocabulary acquisition is very complex and is the subject many books and articles. For that reason, I will only briefly outline some of the relevant details. I will start by summarising what is meant by knowing a word, and then discuss the importance of taking the effects of forgetting into account.

Some teachers, and the vast majority of students, seem to have very simplistic views of what it means to know a word. This usually involves a form-meaning relationship, or more specifically the L2 word and its L1 translation. However this is just the first step, there is a lot more to knowing a word. Nation (2001) gives an extensive list of factors involved in knowing a word, which illustrates just how complicated it actually is:

From the point of view of receptive knowledge and use, knowing the word, for example, *underdeveloped* involves:
• being able to recognise the word when it is heard
• being familiar with its written form so that it is recognised when it is met in reading
• recognising that it is made up of the parts under-, -develop- and -ed and being able to relate these parts to its meaning
• knowing that underdeveloped signals a particular meaning

- knowing what the word means in the particular context in which it has just occurred
- knowing the concept behind the word which will allow understanding in a variety of contexts
- knowing that there are related words like overdeveloped, backward and challenged
- being able to recognise that underdeveloped has been used correctly in the sentence in which it occurs
- being able to recognise that words such as territories and areas are typical collocations
- knowing that underdeveloped is not an uncommon word and is not a pejorative word

From the point of view of productive knowledge and use. knowing the word *underdeveloped* involves:
- being able to say it with correct pronunciation including stress
- being able to write it with correct spelling
- being able to construct it using the right word parts in their appropriate forms
- being able to produce the word to express the meaning underdeveloped
- being able to produce the word in different contexts to express the range of meanings of underdeveloped
- being able to produce synonyms and opposites for underdeveloped
- being able to use the word correctly in an original sentence
- being able to produce words that commonly occur with it
- being able to decide to use or not use the word to suit the degree of formality of the situation (At present developing is more acceptable than underdeveloped which carries a slightly negative meaning.)

<div align="right">Nation, 2001, pp26-28</div>

This list covers productive and receptive knowledge of a word in the three main areas of form, meaning, and use. It is obvious from this that there is a lot more to knowing a word than just being able to give an L1 translation. Learners need to build up a picture of the contexts in which a word is used, or not used; its common collocates; level of politeness; connotations; etc. This is equally the case with formulaic sequences. While it is possible to learn the meaning of a new word after only one exposure, as a rule it takes many exposures. It then takes many more exposures in a variety of contexts to acquire a full knowledge of it. This might suggest that high frequency words will be acquired sooner than low frequency words. However, this may not necessarily be true, because high frequency words tend to have a wider range of meaning and use, so there is much more to learn about them, for example; 'rusty' can be used to describe corroded metal, a skill that has been neglected, or a colour, and it also has the connotation of neglect, whereas 'oxidised' only has the meaning of having undergone a process of chemical reaction involving oxygen. Although both words are synonyms, the higher frequency word places a greater learning burden on students than the low frequency one. This is mitigated somewhat by more frequent exposure to the high frequency word.

Forgetting

As with any process of acquisition, there is also a corresponding process of attrition. One problem that occurs due to the lack of exposure to the target language in many EFL settings is forgetting. It is a fact of life that people forget things. Our brains are highly efficient at processing language, and filtering out items that are not useful. While factors like, difficult spelling or pronunciation, multiple meanings, interferance from similar words, etc. can make a word more difficult to learn, the deciding factor seems to be the number of exposures to the word. The more often a word is encountered, or the stronger the need to use it, the more likely it is to be deemed useful and put into long-term memory, and then subsequent meetings with the word in context help to accumulate more detailed information about its range of forms, meanings and uses. Estimates generally range between 8 and 15 meaningful encounters with a word

before it is remembered. 'Meaningful' means that the word is noticed and that attention is given to it.

Suppose a student is given a list of 100 words to learn for a test at time T. He will learn the words up to time T. If this learning were completely successful, he would get 100% for the test. Let us suppose he did not study these words afterwards, but a surprise test on the exact same words were given at time T+1. We would normally expect some forgetting to have taken place, let us assume that forgetting occurs at a rate of 40% over that period, so the student gets 60% for the test. Then the test is given again after a similar interval at time T+2, and assuming the same rate of forgetting, the student would get 36%, and then at time T+3 he would get 22%, and so on. Figure 2 illustrates this process and the parabolic shape that is formed when these scores are plotted against time is known as the forgetting curve (Pimsleur, 1967; Waring, 2002).
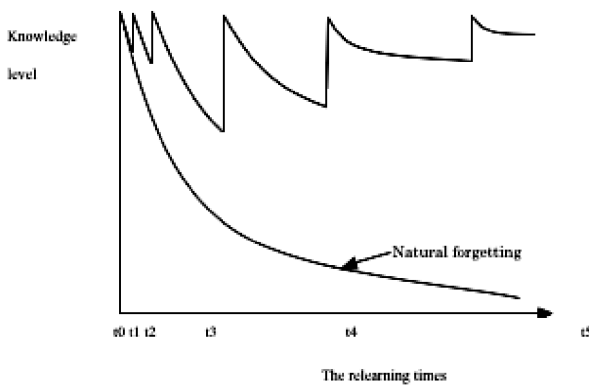

Figure 2: The forgetting curve and a re-learning schedule (Waring, 2002)

This model assumes that there is no practice effect due to the repeated administrations of the test.

The best way to counter the effects of forgetting is to continue studying the words, which generally involves some form of self testing. If this regular revision of the words occurs, then the rate of forgetting should decrease, and the learner can go for increasingly longer periods between practice sessions before a similar number of words

are forgotten. The implications of the forgetting curve are that we cannot assume that if a word or formulaic sequence has been taught, and learned, that it will be remembered. This means that it is important for learners to continually learn and review them in order to acquire a stable vocabulary.

Which words should we teach?

There are a lot of words in any given language, and English has a larger lexicon than most. The New Shorter Oxford English Dictionary on CD-ROM lists 155,637 words; not including categories of words like proper nouns, etc. Some estimates of the total number of words in the English language go as high as 250,000. It is impossible for a native-speaker to know all of these words, let alone an EFL learner. Nation (2001) estimates that an 'educated' native-speaker knows about 20,000 words by age twenty, that means are acquiring approximately 1,000 new words per year. This is still a nearly impossible task for the average EFL learner in Japan who starts learning English at age 12, and does not get anywhere near the same amount of exposure that a native-speaker child would. Therefore, it is important to identify words which will give maximum coverage, and then teach them systematically, but, unfortunately, this often happens in a very haphazard way.

Liu and Nation (1985) found that a learner needs to know at least 95% (one word in twenty) of the words around an unknown word to have a good chance of guessing its meaning from context, but that 98% (one word in fifty) was optimal. This means that a learner needs a vocabulary of about 5,000 word families in order to understand an unmodified text. This is a good long term goal, but it is still rather high for most learners.

Fortunately, corpus linguistics has provided some of the answers to the question of which words to teach. This involves collecting a large body (corpus) of language - usually written, although spoken corpora do exist – and then performing various statistical analyses on it. For the current discussion, frequency data is the most useful. The usual way that this is calculated is to count the number of times that each word occurs, and then rank them from the most to the least frequent. Leech et al (2001) in a

study of the British National Corpus (BNC) of 100-million running words found that of 757,087 different words, 397,041 (52.44%) occur only once. Obviously, such one-off words are of relatively little value to a learner due the low return for the effort required to learn them.

1000 word (lemma) level	% coverage of text (tokens)
1000	72.0
2000	79.7
3000	84.0
4000	86.7
5000	88.6
6000	89.9

Table 1. The percentage text coverage of each successive 1000 lemmas in the Brown Corpus (in Nation, 2001, p15)

Table 1 compares word frequency and coverage. These figures show that gains in coverage decrease rapidly as word frequency decreases. The 2000 most frequent words are generally accepted to be the high frequency words of English, and with nearly 80% coverage of unmodified texts, they offer a good return for learning effort. This means that a sensible approach to teaching vocabulary is to consolidate, and build on, this 2,000 word base. This represents a very useful foundation on which to build learners' vocabulary. Furthermore, there are usually very high correlations between frequency counts based on different corpora up to the 2,000 word level, but beyond that correlations drop due to variations in the composition of the corpora they are based on. We need to obtain similar frequency data for formulaic sequences.

Developing an EFL conversation test

Trying to find good teaching materials for classes is a challenge for even the most experienced teachers. One of the most important considerations in this process is making sure that the level of the materials is not too difficult or too easy. If the materials are too difficult, then no matter how interesting and stimulating they might be, learners will become frustrated and possibly give up. On the other hand, if they are too easy, then learners will not be stimulated, get bored, and, likewise, give up. Most

teachers have experienced the situation where materials that were a huge success in one class, were a complete failure in another.

Textbook selection

English native-speaker teachers working in Japanese universities and schools are frequently asked to teach English conversation classes. There are a wide range of textbooks available from both international and Japanese publishers. In many ways, teachers are spoiled for choice, and it is a difficult task to find one that matches the level of the class, the teacher's general philosophy of language teaching, and teaching style. One problem when selecting a textbook is that most of the "so called" conversation textbooks are just thinly disguised grammar books. Another problem is that many textbooks are too long and it can be very difficult to complete all, or even most of, the book in one semester. In the case of this author, there was a specific set of criteria that a textbook should meet: It should be suitable for low-level learners, if possible, with Japanese support. It should teach the skills and language that learners need in order to participate in conversations in English. Finally, it should teach students how to talk about themselves, as opposed to assuming a persona in order to complete a task based on an imaginary situation that they may or may not have experienced. It took several years of trial and error to find a textbook that met these criteria. One of the key features of the textbook (Ichiyama et al, 2002) is that it breaks a conversation into a series of stages and teaches useful phrases and strategies that can be used in those stages.

The textbook has three units, each centred around a typical situation that students might encounter in their daily lives. Unit 1 is about talking with a person for the first time, unit 2 is about talking with a teacher, and unit 3 is about talking with a friend. Each unit breaks the conversation into 5 stages that follow variations of the basic pattern of greetings, discussion, and ending. For the purposes of this paper, I will focus on unit 1, and the mid-year conversation test I developed based on it. The five stages in unit 1 are entitled; greetings and introductions, small talk, signal to end conversation, ending introductions, and goodbyes. Each stage introduces a series of useful formulaic

sequences typically used in that situation (refer to the grading schedule in the Appendix for a breakdown of these phrases).

Test development

After textbook selection, another important consideration for teachers is assessment. Bachman and Palmer (1996) state that as a basic principle of good test design "[in] order for a language test to be useful for its intended purposes, test performance must correspond in demonstrable ways to language use in non-test situations." In practical terms, this means that a test for an English conversation course should test the ability of learners to participate in a conversation in English. In other words, the final test should involve the students having a conversation.

With the abovementioned principle, in mind, I began the process of finding a conversation test. The problem was how to test students' conversation skills in a practical and objective manner. The textbook's teacher support website has a grading scale for oral interview tests where the teacher acts as the interlocutor. However, it proved to be very difficult to participate in the conversation and to grade impartially at the same time. The next step in the test development process was to test students in pairs. This reduced the overall time required to administer the test, and freed the teacher up to focus on grading. However, the grading scale uses very general descriptors of an ideal performance and five 5-point and 10-point Likert scales, but there are no descriptions of how to grade low or intermediate performances. The grading schedule was difficult to use because scores tended to become more of a general impression of the conversation, which meant that the scores lacked precision and objectivity.

The eventual solution to this problem of precision and objectivity was to develop a checklist (see Appendix) of all of the phrases and strategies from the course textbook laid out in the order that they would typically occur in a conversation. A check mark is placed next to each item as it is used. In the case of items that can be used on multiple occasions, a check mark is placed next to it each time it is used. Each check is counted as 1 point with the exception of the active listening, "Uh-huh" and Mm-Hmm", which get 1 point for the first use and 0.2 points for each subsequent use in order to

discourage padding of scores by their overuse. The conversations are timed and a grade between 1 and 5 points is given based on this time, the range of times for each score is decided based on the overall level of ability of the class. Students are encouraged to speak longer, but in order to discourage stalling for time there is also a fluency grade of 1 to 4 points. Because this textbook and consequently the test are intended for low level learners, students are given time to prepare for the test, and many opt to write their conversations in the form of a script. Student pairs are allowed to bring any support materials that they need to the test, consequently a preparation grade between 1 and 5 points depending on the proportion of the conversation they read with 5 points being awarded to students for a conversation that is completely memorised or spontaneous, and one point for students who read the entire conversation.

This checklist style of grading sheet provides a more objective rating of the students' performances in terms of testing their ability to use the content taught in the course. Furthermore, four or five years of using this test format, have lead to refinements in its layout, use, and scoring, with consequent improvements in its accuracy and objectivity. It also allows the teacher a reasonable degree of flexibility depending on levels of student ability, for example; students can be given more time for preparation in lower level classes, or required to display a greater degree of spontaneity in higher level classes. One ongoing problem has been that the rater tends to be focussed on listening to the conversation and checking off the list, so it is difficult to also monitor visual aspects such as eye contact.

Subjects

The subjects were 483 second-year high-school students studying at a Japanese combined high-school and vocational school specialising in engineering. All subjects were either 16 or 17 years old at the time of testing, and would have completed three years of English study at junior high-school and one and a half years of senior high-school. The data was accumulated from three cohorts of second-year students over a period of three years with each consisting of 160 students divided up into four classes of about 40 students based on their major; mechanical engineering, electrical

engineering, materials engineering, and information technology. All classes were made up predominantly of male students, ranging from about 75% to 100% depending on the major, with materials engineering and information technology tending to have the larger numbers of female students. Though the school places a strong emphasis on English, it is considered a minor subject in respect to the overall curriculum. English is a required subject, and students must attain a grade of 50% or higher in order to pass the course. Each class had three English lessons per week, two with a Japanese teacher of English, and one with a native-speaker teacher. The final grade was assigned by both teachers; 70% by the Japanese teacher and 30% by the native speaker. The native-speaker class is the subject of the current study,

Procedure

The test was administered to determine their mid-year grades, and was administered on completion of Unit 1, which is about talking to a person for the first time. Students were given a copy of the grading schedule during the preparation period, two or three classes prior to administration of the test. Class periods are 50 minutes long. The test procedure was explained to students in Japanese on at least two separate occasions prior to administering the test. The tests were conducted in a small, enclosed office in the classroom while the rest of the class watched a video or prepared for their tests. This was done in order to ensure privacy and to lower affective factors, in particular self-consciousness.

The tests for each class were conducted over two class periods due to the large class sizes and the limited class time available. When each pair entered the test room, they were instructed to sit facing each other, and the test procedure was briefly explained to them again in Japanese. They were allowed a little time to compose themselves. As this test format was new to students, they were understandably nervous, and every effort was made to put them at ease. When the students were ready, they could begin their conversation, and the stopwatch would be started. Students were told that they could bring dictionaries, textbooks, handouts, notes, etc. to the test, but were advised that using them could slow them down and might cost them preparation and

fluency points. They were also told that they could ask the instructor or their partner for help in English if they got stuck. Conversations were in pairs, but if there were an odd number of students in the class, then a student was asked to volunteer to do the test twice. As an incentive to volunteer, the higher of the two scores would be used for their final grade. Students were instructed to do the conversation as if they were meeting their interlocutor for the first time.

Results and Discussion

Because English is not the major subject, the school's unstated policy toward English classes is that a student that is doing well in major subjects should not be held back on account of poor performance in English. As a result, in some cases adjustments needed to be made in order to achieve this. However, the figures quoted below are based on the raw, unadjusted figures.

Time

Students were told that their conversations should be at least one minute, and that this would earn them three points, or 60% of the time grade, and longer conversations earn 4 or 5 points. As can be seen in Figure 3, the majority of students met this requirement or were very close to it. The one-minute minimum requirement accounts for the skewed nature of the distribution. Even though the one-minute
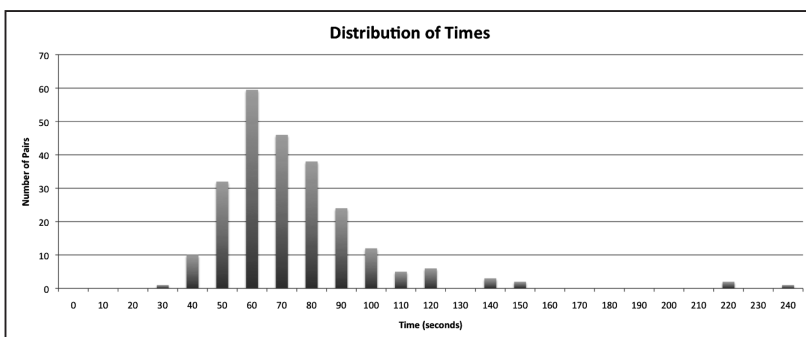

Figure 3: Distribution of pairs' conversation times (seconds)

minimum requirement was explained to students in Japanese at the beginning of every preparation session, as can be seen in Figure 3, a significant number, 43 out of 242 pairs (17.7%), did not achieve this, however, 32 of these were within 10 seconds of it. In some cases these reduced times were due to students speaking more rapidly than in practice, as evidenced by the large number of pairs whose times were between 50 and 59 seconds. In other cases, as the scoring system was explained to students in their L1 and the grading schedule was made available to them, the remaining shorter conversation times can be interpreted as a conscious decision by some pairs to only put in the minimum effort required to obtain a passing grade.

Intuitively, it would seem a reasonable assumption that these reduced times would affect scores in two ways; firstly by reducing the time score, and secondly by reducing the potential for gaining points for other aspects of the conversation, in other words, the longer a pair maintained the conversation, the more they said, and the more opportunities they had to get points. However, this assumption is not completely borne out by the results. There is only a moderate correlation between times and overall scores (Pearson $r = 0.57$). While this figure is significant ($p > .05$), it only accounts for one third ($r^2 = 0.326$) of the variation in scores. These figures remain almost the same when students with extreme outlying times over 160 seconds are eliminated in order to reduce the skewedness of the figures. What this indicates is that while time was a significant factor, other factors such as fluency, and actual use of the formulaic sequences and strategies taught in the course were also significant. For example, generally speaking, four fluency points were awarded to students, but this was reduced to offset the possibility of students obtaining higher time scores by speaking slowly or pausing frequently.

Scores

It must also be emphasised that this test was developed as practical means of student assessment and not as a research instrument, therefore, some of the measurements do not give precise figures for the use of each formulaic sequence.

The primary purpose of any achievement test should be to give students every
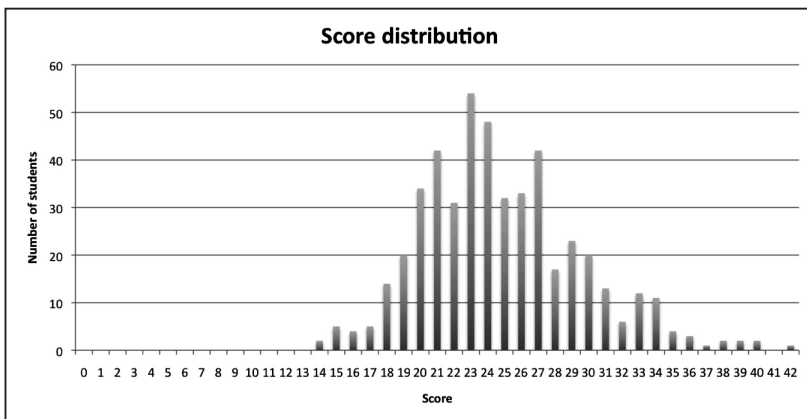
Figure 4: Distribution of scores

opportunity to demonstrate the degree of mastery they have acquired of the course content. Because of the relatively free nature of this conversation test format, it is very important that students have a very clear understanding of what is required of them. As this test format was new to all of the students in these courses, the instructor went to great lengths explain the test format and the grading system. With the exception of two students who both scored 14, all students received passing grades. As can be seen in Figure 4, the scores generally resemble a normal distribution with a mean of 24.81 and a standard deviation of 4.746. In theory there is no maximum score for the test, however, in practice it was found that 30 tended to be a fair cut off. Students achieving scores higher were given the full 30% towards their final course grades, with the remaining 70% being decided by the Japanese teacher. While this would appear to be giving maximum scores to a lot of students, on a class-by-class basis it generally only affects about three or four students. Furthermore, these exceptional students should be seen as having demonstrated greater mastery than required.

Use of conversation strategies and formulaic sequences

As discussed above, there should be a distinction between conversation strategies and formulaic sequences. I will begin by discussing the results of strategy and

pragmatic knowledge of the structure of a conversation. The grading schedule (see Appendix) is laid out in a similar order to that of the textbook, which replicates general order that the stages and their elements should occur. With very few exceptions, every pair of students adhered to this order. This suggests that students did not have trouble with acquiring the pragmatic aspects of the structure of conversations, or that they could successfully transfer their knowledge from L1 to English.

Stage 1 is greetings and introductions. There are three formulaic sequences introduced in this stage; a greeting, introducing yourself, and 'Nice to meet you' and its reply 'Nice to meet you, too'. In addition, there were a few variations that communicate a similar intent, such as; 'Hello' and 'Hi', or 'How are you?' and 'How's it going'. In the test, all students gave some kind of greeting or opening in their conversation, although,

Description	Number of students	Percentage	Average uses per student
Hi! How's it going? Hello. How are you?	455	94.2%	1
My name is… I'm…	459	95.0%	1
Nice to meet you (too)	454	94.0%	1
Total greetings	483	100.0%	2.83

Table 2: Results for Stage 1, Greetings and introductions

as can be seen in Table 2, some did not perform specific parts of the sequence from the textbook with students on average using 2.83 of the three (number of uses ÷ number of users). Of the three parts introduced in this stage, the vast majority of students, 420 (86.97%), used all three, 45 (9.32%) used two, and 18 (3.73%) used only one. In part, this high rate of use is due to coral repetition at the beginning of each lesson to review what had been studied previously. Consequently, because this stage was taught first, it was the one that got the most repetition throughout the semester. In addition, students would have been exposed to them regularly prior to beginning the course.

Stage 2, entitled small talk, primarily focuses on conversation strategies rather

than formulaic sequences. This is due to the wide range of possible formulaic sequences available that would be too difficult to cover adequately in the available time. The strategies covered are; choosing appropriate topics, bringing up new topics using 'So...', developing topics, and active listening using 'uh-huh' and 'mm-hmm'. Choosing appropriate topics refers to avoiding potentially embarrassing topics, and the textbook lists six questions that are often used in conversations when meeting a person for the first time. Developing topics includes; adding extra information when answering questions (A+ answers), and asking a follow up question or making a comment about what the interlocutor has just said.

The results for raising new topics in stage 2, shown in Table 3, are mixed. Of the six questions suggested as appropriate topics, five were only used by between 0.6% and 6.6% of students. Only 'Where are you from?' was used by a fairly large number of students (57.1%). This appears to be due to some of the questions being inappropriate for the students' situation, for example, students studying the same major at the same school would have had trouble developing this as a topic. Many of the students who

Description	Number of students	Percentage	Average uses per student
1. Appropriate topics			
Where are you from?	276	57.1%	1
What do you do?	32	6.6%	1
What's your major?	13	2.7%	1
Is this your first time in…?	3	0.6%	1
How do you like…?	25	5.2%	1
How long have you been in…?	11	2.3%	1
Made their own	262	54.2%	1.36
Total questions	**409**	**84.7%**	**1.17**
2. Bring up new topics			
So…	298	61.7%	1.39

Table 3: Results for Stage 2, Small talk – raising topics

used these questions had adopted alternate personas for the test. 'Where are you from?' was used in the model conversation, which many students referred to during preparation, and so may have been more strongly reinforced than the others. Many students preferred to make their own questions, and most asked follow up questions. The 'Total questions' in Table 3 is the number of students who used at least one of the sample questions, or made their own questions. Most students (84.7%) asked at least one question of some kind to initiate a topic, the remainder having opted for a more one-sided, interview style conversation.

Until taking this course most students did not appear to have encountered 'so' as discourse marker introducing a new topic. Therefore, the instructor made a point of emphasising it and made a point of putting it before questions during choral drills at the beginning of class. Of the students who used it, 94 used it more than once, giving the relatively high rate 1.39 times per user. While use by two thirds of students is encouraging, it could be better. In the first two years of administering the test, one point was given for the first use and 0.2 of a point for each subsequent use to discourage overuse, however, in the third year, this policy was changed to one point for every appropriate use to encourage proper use. This change did not affect the number of users (94, 106, and 98 respectively) or their rate of use (1.36, 1.39, and 1.40 respectively).

Developing topics was very strongly emphasised in class, and students were told

Description	Number of students	Percentage	Average uses per student
3. Developing Topics			
A+E answers	412	85.3%	1.93
Follow up questions	354	73.3%	1.66
Comments	328	67.9%	1.47
Total topic development	**477**	**98.8%**	**3.91**
4. Active Listening			
Uh-huh, Mm-hmm	132	27.3%	1.05

Table 4: Results for Stage 2, Small talk – developing topics

at every opportunity that it was the most important strategy in the book, and that it was a way that they could gain extra points. This emphasis appears to have paid off (see Table 4), because, with only a handful of exceptions, almost all students used at least one of the strategies for developing topics. Also, the average number of times that each student developed a topic is fairly good, however, the figures for each of the three strategies are low in comparison with adding extra information to answers (A+ answers) being the most common followed by follow up questions and comments, respectively.

The rate of active listening as evidenced by use of 'uh-huh' and 'mm-hmm' was disappointing. Many of the students who used it displayed a degree of discomfort or amusement when using it, and this may account for the low rate of use. It could be that its nonverbal nature and the importance of timing, which is different to the Japanese equivalent, were also contributing factors.

An important reason for choosing the textbook was that it is one of the few available that teaches how to end a conversation politely, in particular, how to tell an interlocutor that it is time to end the conversation. This lack of prior exposure is reflected in the relatively low rate of use of formulaic sequences from stage 3 compared to the very familiar ones from stage 5 and the slightly less familiar ones from stage 4. As stage 3 is taught later in the course, students do not appear to be getting sufficient

Description	Number of students	Percentage	Average uses per student
Stage 3 – Signal to end conversation			
I'm sorry, but I have to go.	276	57.1%	1
I have to…	235	48.7%	1
Stage 4 – Ending introductions			
It was nice meeting / talking to you (too)	412	85.3%	1
Stage 5 – Goodbyes			
See you/ya, Goodbye.	472	97.7%	1

Table 5: Ending the conversation

repetition during choral drills to allow them to gain confidence in using them. This needs to be remedied in future courses.

Conclusion

This course and test developed from a belief that content and assessment should reflect the goals of the course. This development process incorporates principles about the content of courses, the way that they are taught, and a general philosophy of language and language teaching. It appears that a significant proportion of the language that we produce is formulaic. I think that conversation strategies and formulaic sequences as vocabulary items, offer exciting possibilities for the teaching and learning of discourse. This seems to be particularly true of speaking where the processing demands are very high. Future research needs continue to develop our knowledge of formulaic sequences, and investigate the kind of formulaic sequences that native speakers use, what we should teach students about them, and how we can assess them. Of particular importance is developing frequency-based pedagogical lists of formulaic sequences.

References

Bachman, L. F. & Palmer, A. S. (1996). *Language Testing in Practice.* UK: Oxford University Press.

Ichiyama, M. S., Critchley, M. P., Fiyouzat, R., Iijima, Y., & Suzuki, T. A. (2002) *Encounters: A new approach to English Communication, Book 1.*Tokyo: Nan'Un-Do Publishing Co., Ltd.

Kenny, T. and Woo, L. (2004). *Nice Talking with You – Second edition, Teacher's Manual.* Tokyo: MacMillan LanguageHouse

Leech, G., Rayson, P., and Wilson, A. (2001). *Word frequencies in written and spoken English.* Great Britain: Longman

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge MA: Bradford

Lewis, M. (1993). *The Lexical Approach.* Boston MA: Heinle.

Liu Na and Nation, I. S. P. (1985) "Factors affecting guessing vocabulary in context". *RELC Journal.* 16; 33-42.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language.* Great Britain: Cambridge University Press.

Pimsleur, P. 1967. "A memory schedule". *Modern Language Journal.* 51(2): 73-75.

Read, J, (2000), *Assessing Vocabulary,* Cambridge University Press, Great Britain.

Schmitt, N. (ed.). (2004). *Formulaic Sequences.* Amsterdam: John Benjamins Publishing Company.

Waring, R. (2002). *In defence of learning words in word pairs - but only when doing it the 'right' way!* Retrieved 10 Oct. 2003 from

http://www1.harenet.ne.jp/~waring/vocab/principles/systematic_learning.htm

Wray, A. (2002). *Formulaic Language and the Lexicon.* Cambridge, UK: Cambridge University Press.

Appendix

Grading Check Sheet

Class _____

_____   **Names**   _____

- **Stage 1 – Greetings and Introductions**
  - ▪ Hi! How's it going?   Hello. How are you?
  - ▪ My name is… I'm…
  - ▪ Nice to meet you (too)
  - ▪ EYE CONTACT

- **Stage 2 – Small Talk**
  - **1. Appropriate topics**
    - • Questions p.10
      - o Where are you from?
      - o What do you do?
      - o What's your major?
      - o Is this your first time in…?
      - o How do you like…
      - o How long have you been in…?
    - • Made their own
  - **2. Bring up new topics**
    - • So…
  - **3. Develop Topics**
    - • A+E answers
    - • Follow up questions
    - • Comments about things said
  - **4. Active Listening**
    - • Uh-huh, Mm-hmm

- **Stage 3 – Signal to end conversation**
  - ▪ I'm sorry, but I have to go.
  - ▪ I have to…

- **Stage 4 – Ending introductions**
  - ▪ It was nice meeting/talking to you (too)

- **Stage 5 – Goodbyes**
  - ▪ See you/ya, Goodbye.

General Comments
- **Fluency**
  - o no significant stops and starts
  - o few significant stops and starts
  - o some stops and starts
  - o many stops and starts

- **Preparation**
  - o No notes
  - o Glancing
  - o Read occasionally
  - o Mostly reading
  - o Reading

Unit 1 Grading schedule

Time_____:_____